

REFERENCE ONLY



## UNIVERSITY OF LONDON THESIS

Degree phd

Year 2007

Name of Author ANGHARAD PUW  
DAVIES

### COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

### COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

### LOAN

Theses may not be lent to individuals, but the University Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: The Theses Section, University of London Library, Senate House, Malet Street, London WC1E 7HU.

### REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the University of London Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The University Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

***This thesis comes within category D.***

☐

This copy has been deposited in the Library of \_\_\_\_\_

☐

This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.



**A STUDY OF RESUSCITATION-PROMOTING FACTORS IN**  
***MYCOBACTERIUM TUBERCULOSIS***

**Ph.D. thesis**

**submitted by Angharad Puw Davies**

**Royal Free and University College Medical School**

**2006**

UMI Number: U592782

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U592782

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346



## Abstract

Tuberculosis is a major threat to human health. About one third of the world's population is latently infected with *Mycobacterium tuberculosis*. In these cases the bacillus is in a state of low metabolic activity, making eradication difficult with conventional chemotherapy, which targets actively metabolizing organisms. The mechanisms by which *M. tuberculosis* reactivates to cause disease are currently unknown but a better understanding could greatly improve the treatment of tuberculosis. Resuscitation-promoting factor is a protein first identified in the supernatant of stationary phase cultures of *Micrococcus luteus*. It is active in picomolar concentrations, increasing the number of culturable *M. luteus* cells from dormant populations and shortening the lag phase of growth of small inocula. Bioinformatic searches reveal over 40 examples of *rpf*-like genes in the high G+C cohort of Gram-positive bacteria, including *M. tuberculosis*, which contains five *rpf* gene orthologues. The work presented here investigated aspects of the *M. tuberculosis* Rpf's. Improvements in solubility of recombinant mycobacterial (*M. tuberculosis* and *M. smegmatis*) Rpf's were achieved by manipulating induction times and temperatures during protein expression and by using new hosts and vectors and producing novel fusion proteins. New assays were devised to measure the biological activity of recombinant Rpf's, using ATP bioluminescence of *M. luteus* cultures. A phage display library for *M. tuberculosis* was constructed, in an attempt to identify a protein receptor for Rpf. Rpf expression in human infection was investigated for the first time, using immunocytochemistry. Anti-Rpf antibodies were applied to human tissue sections infected with *M. tuberculosis*. Rpf was found to be located within epithelioid giant cells and in the immediate vicinity of acid-fast bacilli in necrotic centres. The presence of Rpf in human tuberculosis infection demonstrated in this work suggests that Rpf's may have a role in controlling dormancy of the bacilli in human disease.

## CONTENTS

	Page
<b>ABSTRACT</b>	<b>2</b>
<b>CONTENTS</b>	<b>3</b>
<b>LIST OF FIGURES</b>	<b>8</b>
<b>LIST OF TABLES</b>	<b>13</b>
<b>LIST OF ABBREVIATIONS</b>	<b>14</b>
<b>ACKNOWLEDGEMENTS</b>	<b>15</b>
<b>CHAPTER I                      INTRODUCTION</b>	<b>17</b>
<b>1.1            Tuberculosis and latency – the clinical problem</b>	<b>18</b>
<b>1.2            Dormancy</b>	<b>20</b>
1.2.1      Definition	20
1.2.2      Site and physiology of dormant organisms	21
1.2.3      Molecular changes in dormancy	23
<b>1.3            Resuscitation-promoting factors (Rpf)</b>	<b>24</b>
1.3.1      Discovery of Rpf	25
1.3.2      Structure of Rpf and possible mode of action	26
1.3.3      Activity of Rpf	29
1.3.4      Onset of transcription and appearance of Rpf in cultures	30
1.3.5      Rpfs in <i>Mycobacterium tuberculosis</i>	31
1.3.6      Effect of Rpf expression <i>in vivo</i>	35
1.3.7      Clinical relevance of Rpf	36
<b>1.4            Aims of the project</b>	<b>37</b>
<b>CHAPTER II                    MATERIALS AND METHODS</b>	<b>38</b>
<b>2.1            Cloning experiments</b>	<b>39</b>
2.1.1      DNA extractions from <i>Escherichia coli</i>	39
2.1.2      DNA extractions from <i>M. tuberculosis</i>	39
2.1.3      Measurement of DNA concentrations	40
2.1.4      PCR for genes <i>rpjD</i> and <i>Rv0328</i>	40
2.1.5      TOPO 2.1® cloning reaction	41
2.1.6      Restriction endonuclease digestions	41
2.1.7      Agarose gel electrophoresis	41
2.1.8      Extraction of DNA from agarose gel	42
2.1.9      Treatment with calf-intestinal alkaline phosphatase	42
2.1.10     Ligation reactions	42
2.1.11     Preparation of competent <i>E. coli</i> host cells	43
2.1.12     Transformation of competent cells	43
2.1.13     Blue-white selection of XL2-Blue strain	43
2.1.14     Sequencing of cloned gene inserts	44
2.1.15     Storage of expression hosts and clones	44
<b>2.2            Protein expression</b>	<b>44</b>
2.2.1      IPTG induction	44
2.2.2      Harvesting the induced cells	45
<b>2.3            Protein extraction and purification</b>	<b>45</b>
2.3.1      Sonication	45
2.3.2      Protein purification by metal-ion chelation column chromatography	46
2.3.2.1    Protein purification using in-house sepharose columns	46

2.3.2.2	Protein purification using Quick-900 columns	47
2.3.3	Estimation of purified protein concentration	47
2.3.4	Large scale expression, extraction and purification of RpfE	48
2.3.5	TEV protease cleavage of His-tag from RpfE expressed from pET-15b. <i>tev</i>	49
2.3.6	SDS-PAGE (Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis)	49
2.3.7	Immuno-blotting	50
<b>2.4</b>	<b>Establishing dormant cultures of <i>M. bovis</i> BCG</b>	<b>51</b>
2.4.1	Tubes	52
2.4.1.1	Tubes - culture and sampling	52
2.4.1.2	Tubes - estimation of viable count	52
2.4.2	Flasks	52
2.4.2.1	Flasks - culture and sampling	52
2.4.2.2	Flasks - estimation of viable count	53
2.4.2.3	Flasks - estimation of Most Probable Number	53
<b>2.5</b>	<b>Protein activity testing</b>	<b>53</b>
2.5.1	Protein activity testing measuring optical density of <i>Micrococcus luteus</i> cultures	53
2.5.2	Protein activity testing using <i>M. bovis</i> BCG cultures	54
2.5.2.1	'Concentration-activity test' method	55
2.5.2.2	Most Probable Number method	56
2.5.3	Protein activity testing using ATP bioluminescence of <i>M. luteus</i> cultures	57
2.5.3.1	Overnight <i>M. luteus</i> cultures ('overnight test')	57
2.5.3.2	Four hour <i>M. luteus</i> cultures ('rapid test')	58
<b>2.6</b>	<b>Phage display</b>	<b>58</b>
2.6.1	Construction of phage display library	58
2.6.2	Transfection and PEG precipitation of phage	60
2.6.3	Titering of phage stocks and fractions from panning	60
2.6.4	Panning for RpfB receptor protein	61
2.6.5	Sequencing and analysis of inserts from clones selected by panning	62
2.6.6	Cloning and expression of Rv0328	62
<b>2.7</b>	<b>Immunocytochemistry of <i>M. tuberculosis</i>-infected tissues with anti-Rpf antibodies</b>	<b>62</b>
2.7.1	Immunocytochemistry protocol	63
2.7.2	Absorption of anti-Rpf antibodies using RpfB	64
2.7.3	Ziehl-Neelson staining of sections	64
2.7.4	Staining of sections with anti-lipoarabinomannan antibodies	65
<b>CHAPTER III</b>	<b>RESULTS: CLONING AND EXPRESSION OF <i>M. TUBERCULOSIS</i> Rpfs</b>	<b>66</b>
<b>3.1</b>	<b>Introduction</b>	<b>67</b>
3.1.1	Aims	67
3.1.2	Host strain	67
3.1.3	Vectors	68
<b>3.2</b>	<b>Materials and Methods</b>	<b>69</b>
3.2.1	Cloning of <i>rpfD</i> into pET-21a vector and Origami B (DE3)pLysS host	69
3.2.2	Transformation of Origami B (DE3) pLysS host cells with <i>rpfA</i> , <i>rpfB</i> , <i>rpfC</i> and <i>rpfE</i> .	71
3.2.3	Expression and purification of the five <i>M. tuberculosis</i> Rpfs	71
<b>3.3</b>	<b>Results</b>	<b>75</b>
3.3.1	Expression and purification of RpfA	75
3.3.2	Expression and purification of RpfB	75
3.3.3	Expression and purification of RpfC	76
3.3.4	Expression and purification of RpfD	77
3.3.5	Expression and purification of RpfE	78
<b>3.4</b>	<b>Discussion</b>	<b>79</b>

<b>CHAPTER IV</b>	<b>RESULTS: IMPROVING SOLUBILITY AND EXPRESSION OF RpfS</b>	<b>81</b>
<b>4.1</b>	<b>Introduction</b>	<b>82</b>
4.1.1	Recombinant protein expression in <i>E. coli</i>	82
4.1.2	Host strain	83
4.1.3	Fusion proteins	84
4.1.4	Induction temperature	84
4.1.5	Aims	85
<b>4.2</b>	<b>Improving the solubility and yield of Rpf from <i>Mycobacterium smegmatis</i> by using a new expression host and a fusion protein</b>	<b>86</b>
4.2.1	<i>M. smegmatis</i> Rpf (Rpf <sub>smeg</sub> )	86
4.2.2	Materials and Methods	87
4.2.2.1	Subcloning of Rpf <sub>smeg</sub> into NusA fusion protein and Origami (DE3) host	87
4.2.2.2	Expression of Rpf <sub>smeg</sub> with <i>M. luteus</i> Rpf for comparison	92
4.2.2.3	Immunoblotting for Rpf <sub>smeg</sub> and <i>M. luteus</i> Rpf	92
4.2.2.4	Extraction and purification of Rpf <sub>smeg</sub>	92
4.2.3	Results	93
4.2.3.1	Expression of Rpf <sub>smeg</sub>	93
4.2.3.2	Extraction and purification of Rpf <sub>smeg</sub>	96
<b>4.3</b>	<b>Improving the yield and solubility of RpfB and RpfE by changing the expression host and induction temperatures, and scaling up of production to achieve greater yields</b>	<b>99</b>
4.3.1	Materials and Methods	100
4.3.1.1	Generating clones of <i>rpfB</i> and <i>rpfE</i> in a new host and vector to improve solubility	100
4.3.1.2	Expression of RpfB and RpfE from clones AD5, AD8 and AD11	100
4.3.1.3	Large scale expression of RpfE from clone AD11	101
4.3.2	Results	101
4.3.2.1	Expression of RpfB from clone AD5 ( <i>rpfB</i> in pET-19b vector and Origami B (DE3) pLysS host	101
4.3.2.2	Expression of RpfE from clone AD8 ( <i>rpfE</i> in pET-19b vector and Origami B (DE3) pLysS host	102
4.3.2.3	Expression of RpfE from clone AD11	102
4.3.2.4	Scaling up of production of RpfE	103
<b>4.4</b>	<b>Discussion</b>	<b>111</b>
<b>CHAPTER V</b>	<b>RESULTS: ACTIVITY TESTING OF RPF USING CULTURES OF <i>MYCOBACTERIUM BOVIS</i> BCG AND <i>MICROCOCCUS LUTEUS</i></b>	<b>113</b>
<b>5.1</b>	<b>Introduction</b>	<b>114</b>
<b>5.2</b>	<b><i>In vitro</i> models of dormancy in <i>M. tuberculosis</i></b>	<b>115</b>
<b>5.3</b>	<b>Most Probable Number tests</b>	<b>117</b>
<b>5.4</b>	<b>Establishing a dormancy model in <i>M. bovis</i> BCG</b>	<b>117</b>
5.4.1	Method: setting up of cultures	117
5.4.2	Results	118
5.4.2.1	Tubes - viable counts	118
5.4.2.2	Flasks - viable counts	119
5.4.3	Discussion - dormancy model in <i>M. bovis</i> BCG	120
<b>5.5</b>	<b>Activity testing of mycobacterial RpfS by MPN and concentration-activity tests on <i>M. bovis</i> BCG</b>	<b>122</b>
5.5.1	Methods	123
5.5.1.1	MPN test	123
5.5.1.2	'Concentration-activity' test	123
5.5.2	Results	123
5.5.2.1	Results RpfA: concentration-activity test	123
5.5.2.2	Results RpfB: MPN test	125
5.5.2.3	Results RpfC: MPN test	126

5.5.2.4	Results RpfD: concentration-activity test	127
5.5.2.5	Results RpfE: concentration-activity test	127
5.5.2.6	Results Rpf <sub>smeg</sub> : concentration-activity tests and MPN tests	127
5.5.2.6.1	Experiment 1: concentration-activity test	128
5.5.2.6.2	Experiment 2: MPN test	129
5.5.2.6.3	Experiment 3: MPN and concentration-activity tests	130
5.5.2.6.4	Experiment 4: MPN and concentration-activity tests	130
5.5.3	Summary of results of activity testing of mycobacterial RpfS on dormant cultures of <i>M. bovis</i> BCG	132
<b>5.6</b>	<b>Activity testing by measuring OD<sub>600</sub> of <i>M. luteus</i> cultures in microtitre plates</b>	<b>133</b>
<b>5.7</b>	<b>Activity testing using ATP bioluminescence of <i>M. luteus</i> cultures</b>	<b>135</b>
5.7.1	Overnight ATP bioluminescence activity test	136
5.7.1.1	Method	136
5.7.1.2	Results	137
5.7.1.2.1	Results: positive control	137
5.7.1.2.2	Results: RpfB from clone AD5	139
5.7.1.2.3	Results: RpfB from clone CA1002	142
5.7.1.3	Conclusions	144
5.7.2	Rapid ATP bioluminescence activity test	145
5.7.2.1	Method	145
5.7.2.2	Results	145
5.7.2.2.1	ATP levels over 76 hours	145
5.7.2.2.2	OD <sub>600</sub> and protein levels over 76 hours	146
5.7.2.2.3	Limiting factors for growth	149
5.7.2.3	Conclusions	151
<b>5.8</b>	<b>Activity testing of Rpf – summary and discussion</b>	<b>151</b>
<b>CHAPTER VI</b>	<b>RESULTS: USE OF PHAGE DISPLAY TO IDENTIFY POSSIBLE CANDIDATES FOR AN RPF RECEPTOR</b>	<b>154</b>
<b>6.1</b>	<b>Introduction</b>	<b>155</b>
6.1.1	Phage display	155
6.1.1.1	Phage display vehicles	156
6.1.1.2	Library amplification	157
<b>6.2</b>	<b>Construction of phage display library</b>	<b>158</b>
<b>6.3</b>	<b>Transfection of phage, PEG precipitation of phage and titering of phage stocks</b>	<b>162</b>
<b>6.4</b>	<b>Panning for RpfB receptor protein</b>	<b>162</b>
6.4.1	First round of panning	163
6.4.2	Second round of panning	164
6.4.3	Third round of panning	165
<b>6.5</b>	<b>Sequencing and analysis of inserts from clones selected by panning</b>	<b>166</b>
<b>6.6</b>	<b>Cloning and expression of Rv0328</b>	<b>172</b>
<b>6.7</b>	<b>Discussion</b>	<b>174</b>
<b>CHAPTER VII</b>	<b>RESULTS: IMMUNOCYTOCHEMISTRY USING ANTI-RPF ANTIBODIES</b>	<b>179</b>
<b>7.1</b>	<b>Introduction</b>	<b>180</b>
7.1.1	Aim	180
7.1.2	Histobacteriology of tuberculosis	180
7.1.3	Location of <i>M. tuberculosis</i> in tissue sections	183
<b>7.2</b>	<b>Method</b>	<b>185</b>
7.2.1	Determining the correct dilution of secondary antibodies	185
7.2.2	Determining the optimal dilution of anti-Rpf antibodies, and immuno-staining of tissue sections	186

7.2.3	Absorption experiment using recombinant RpfB	186
7.2.4	Staining of giant cells in sections not infected with <i>M. tuberculosis</i>	187
7.2.5	Ziehl-Neelson staining of sections	187
7.2.6	Staining with anti-LAM antibodies	187
<b>7.3</b>	<b>Results</b>	<b>188</b>
7.3.1	Staining of <i>M. tuberculosis</i> -infected human tissue with anti-Rpf antibodies	188
7.3.2	Absorption experiment with recombinant Rpf	188
7.3.3	Staining of giant cells in sections not infected with <i>M. tuberculosis</i>	194
7.3.4	Ziehl-Neelson staining of sections	194
7.3.5	Staining with anti-LAM antibodies	194
<b>7.4</b>	<b>Discussion</b>	<b>198</b>
<b>CHAPTER VIII DISCUSSION</b>		<b>201</b>
8.1	Discussion	202
8.2	Further Work	209
<b>REFERENCES</b>		<b>213</b>
<b>APPENDICES</b>		<b>227</b>
Appendix i	DNA sequences, amino acid sequences and composition analysis of <i>M. tuberculosis</i> RpfB	228
Appendix ii	Media and solutions	239
Appendix iii	Table of clones used	246
Appendix iv	Vector maps	247
Appendix v	MPN table	253
Appendix vi	Table showing viable counts in <i>M. bovis</i> BCG dormancy experiments (flask model) on solid agar and liquid media with and without Rpf	254
Appendix vii	From luciferase to Lazarus – winning essay MRC Max Perutz prize 2003	255

	<b>List of Figures</b>	<b>Page</b>
<b>Figure 1.1</b>	Conserved regions of <i>M. tuberculosis</i> Rpf s	27
<b>Figure 1.2</b>	Diagrammatic representation of <i>M. tuberculosis</i> Rpf s	32
<b>Figure 3.1</b>	Flowchart of procedure for Rpf induction and expression	73
<b>Figure 3.2</b>	Flowchart of procedure for Rpf extraction and purification	74
<b>Figure 3.3</b>	Purified <i>M. tuberculosis</i> Rpf A – SDS-PAGE	75
<b>Figure 3.4</b>	Purified <i>M. tuberculosis</i> Rpf B – SDS-PAGE	76
<b>Figure 3.5</b>	Purified <i>M. tuberculosis</i> Rpf C – SDS-PAGE	77
<b>Figure 3.6</b>	Purified <i>M. tuberculosis</i> Rpf D – SDS-PAGE	78
<b>Figure 3.7</b>	Purified <i>M. tuberculosis</i> Rpf E – SDS-PAGE	78
<b>Figure 3.8</b>	All five recombinant <i>M. tuberculosis</i> Rpf s – SDS-PAGE	79
<b>Figure 4.1</b>	Diagrammatic representation of vector pET43.1b (Novagen)	85
<b>Figure 4.2.1</b>	Agarose gel electrophoresis of digested plasmid pSDY5 containing Rpf <sub>smeg</sub> insert	88
<b>Figure 4.2.2</b>	Agarose gel electrophoresis of digested plasmid pSDY5 and vector pET-43.1b	89
<b>Figure 4.2.3</b>	Agarose gel electrophoresis of putative clones containing Rpf <sub>smeg</sub>	90
<b>Figure 4.2.4</b>	2% agarose gel electrophoresis of DNA from Origami (DE3) transformants encoding Rpf <sub>smeg</sub>	91
<b>Figure 4.2.5</b>	0.8% agarose gel electrophoresis of DNA from Origami (DE3) transformants encoding Rpf <sub>smeg</sub>	91
<b>Figure 4.2.6</b>	SDS-PAGE of crude extracts of recombinant <i>M. luteus</i> Rpf and Rpf <sub>smeg</sub>	93

<b>Figure 4.2.7</b>	Immunoblot of crude extracts of recombinant <i>M. luteus</i> Rpf and Rpf <sub>smeg</sub>	95
<b>Figure 4.2.8</b>	SDS-PAGE of recombinant Rpf <sub>smeg</sub> extraction fractions – before optimisation of expression	97
<b>Figure 4.2.9</b>	Immunoblot of recombinant Rpf <sub>smeg</sub> extraction fractions – before optimisation of expression	97
<b>Figure 4.2.10</b>	SDS-PAGE of purified recombinant Rpf <sub>smeg</sub>	98
<b>Figure 4.3.1a</b>	SDS-PAGE of eluted protein from AD11 after HPLC	105
<b>Figure 4.3.1b</b>	SDS-PAGE of eluted protein from AD11 after HPLC	105
<b>Figure 4.3.2</b>	Chromatogram of desalting of protein fractions from AD11 (HPLC)	106
<b>Figure 4.3.3</b>	Chromatogram of desalting of protein fractions from AD11 (HPLC)	107
<b>Figure 4.3.4</b>	Immunoblot of RpfE after purification by HPLC, with and without TEV protease digest	109
<b>Figure 4.3.5</b>	Chromatogram of separation of RpfE from TEV protease (HPLC)	110
<b>Figure 5.1</b>	Viable counts of <i>M. bovis</i> BCG in Sauton's medium, in stationary tubes, over time	118
<b>Figure 5.2</b>	Viable counts of <i>M. bovis</i> BCG in Sauton's medium, in agitated flasks, over time	120
<b>Figure 5.3</b>	Growth curve of <i>M. bovis</i> BCG with and without RpfB	125
<b>Figure 5.4</b>	Growth curve of <i>M. bovis</i> BCG with and without RpfC	126
<b>Figure 5.5</b>	Growth curve of <i>M. bovis</i> BCG with and without <i>M. smegmatis</i> Rpf	130
<b>Figure 5.6</b>	ATP bioluminescence activity testing – positive control	138



<b>Figure 5.7</b>	ATP bioluminescence activity testing – positive control and 12nM 138 RpfB from clone CA1002	138
<b>Figure 5.8</b>	ATP bioluminescence activity test – 100nM RpfB from clone AD5	140
<b>Figure 5.9</b>	ATP bioluminescence activity test – 60nM RpfB from clone AD5	140
<b>Figure 5.10</b>	ATP bioluminescence activity test – range of RpfB concentrations	141
<b>Figure 5.11</b>	ATP bioluminescence activity test (RpfB) read at 12 hours	141
<b>Figure 5.12</b>	ATP bioluminescence activity test (RpfB) read at 16 hours	142
<b>Figure 5.13</b>	ATP bioluminescence activity test–1nM RpfB from clone CA1002	143
<b>Figure 5.14</b>	ATP concentration in <i>M. luteus</i> cultures with and without RpfB over 76 hours	146
<b>Figure 5.15a</b>	Protein content of <i>M. luteus</i> culture + - RpfB over 76 hours	148
<b>Figure 5.15b</b>	Protein content of <i>M. luteus</i> culture +/- RpfB over 24 hours	148
<b>Figure 5.16</b>	OD <sub>600</sub> content of <i>M. luteus</i> culture + - RpfB over 76 hours	149
<b>Figure 5.17</b>	Addition of carbon source to <i>M. luteus</i> stationary phase cultures with and without RpfB	150
<b>Figure 6.1</b>	Agarose gel electrophoresis of <i>M. tuberculosis</i> DNA restricted with blunt-ended enzymes	159
<b>Figure 6.2</b>	Agarose gel electrophoresis of <i>M. tuberculosis</i> DNA restricted with blunt-ended enzymes – reaction scaled up	160
<b>Figure 6.3a-c</b>	Agarose gel electrophoresis of restricted DNA of 20 clones tested during construction of <i>M. tuberculosis</i> phage display library	161
<b>Figure 6.4</b>	Agarose gel electrophoresis of 12 clones after three rounds of panning of phage display library against RpfB	166

<b>Figure 6.5</b>	Consensus sequences of four sequenced inserts from clones isolated 168 after three rounds of panning against RpfB	168
<b>Figure 6.6</b>	Sequence alignment of two near-identical inserts from clones isolated after three rounds of panning	169
<b>Figure 6.7</b>	Diagram of position of Rv0328 within <i>M. tuberculosis</i> genome	170
<b>Figure 6.8a</b>	Gene sequence of <i>M. tuberculosis</i> Rv0328	171
<b>Figure 6.8b</b>	Protein sequence of <i>M. tuberculosis</i> Rv0328	171
<b>Figure 6.9</b>	Agarose gel electrophoresis of Rv0328 amplified by PCR from <i>M. tuberculosis</i> genomic DNA	172
<b>Figure 6.10</b>	Agarose gel electrophoresis of restricted clones containing Rv0328 inserts	173
<b>Figure 7.1</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, unstained negative control, x400 magnification.	190
<b>Figure 7.2a</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, unstained negative control, x1000 magnification.	191
<b>Figure 7.2b</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, immunostained with anti-Rpf antibodies, x1000.	191
<b>Figure 7.2c</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, immunostained with anti-Rpf antibodies, x1000.	191
<b>Figure 7.3a</b>	Lymphoid follicle in section of <i>M. tuberculosis</i> -infected human gut lymph node, negative control, x1000.	192
<b>Figure 7.3b</b>	Lymphoid follicle in section of <i>M. tuberculosis</i> -infected human gut lymph node, immuno-stained with anti-Rpf antibodies, x1000.	192
<b>Figure 7.4a</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, blocked with Rpf prior to immunostaining with anti-Rpf antibodies, x1000.	193

<b>Figure 7.4b</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, unblocked, immunostained with anti-Rpf antibodies, x1000.	193
<b>Figure 7.5a</b>	Section of embolised human liver containing giant cells, unstained, x1000	195
<b>Figure 7.5b</b>	Section of embolised human liver containing giant cells, immunostained with anti-Rpf antibodies, x1000	195
<b>Figure 7.6a-b</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, stained by Ziehl-Nielson and anti-Rpf, x1000.	196
<b>Figure 7.7a</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, immunostained with anti-LAM antibodies, x1000.	197
<b>Figure 7.7b</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, stained by Ziehl-Nielson and anti-Rpf, x1000.	197
<b>Figure 7.7c</b>	Section of <i>M. tuberculosis</i> -infected human gut lymph node, immunostained with anti-LAM antibodies, x1000.	197
<b>Figure 7.8</b>	Diagrammatic representation of distribution of immuno-staining with anti-Rpf antibody within a granuloma.	199

	<b>List of Tables</b>	<b>Page</b>
<b>Table 5.1</b>	Result of RpfA concentration-activity test at 9 days	124
<b>Table 5.2</b>	Result of RpfD concentration-activity test at 10 days	127
<b>Table 5.3</b>	Result of Rpf <sub>smeg</sub> concentration-activity test (experiment 1) at 15 days	129
<b>Table 5.4</b>	Result of Rpf <sub>smeg</sub> concentration-activity test (experiment 4) at 7 days	131
<b>Table 5.5</b>	Summary of results of Rpf <sub>smeg</sub> activity testing	132
<b>Table 5.6</b>	Summary of results of activity testing using <i>M. luteus</i> in microtitre plates	135
<b>Table 5.7</b>	Summary of results of overnight ATP bioluminescence activity tests with RpfB	144
<b>Table 6.1</b>	Results of titration of phage stocks of phage display library	162
<b>Table 6.2</b>	Phage titres after first round of panning of phage display library (all fractions)	163
<b>Table 6.3</b>	Phage titres after first round of panning of phage display library, after amplification	164
<b>Table 6.4</b>	Phage titres after second round of panning of phage display library (bound fraction)	164
<b>Table 6.5</b>	Phage titres after third round of panning of phage display library (bound fraction)	165
<b>Table 7.1</b>	Determination of correct dilution of secondary antibody for immuno-staining	185

## List of Abbreviations

aa	amino acids
ADC	albumin-dextrose complex
bp	base pairs
BCG	Bacille Calmette – Guerin
C	Celsius
cm	centimeter
DNA	deoxyribonucleic acid
ELISA	enzyme linked immunosorbent assay
fM	femtomolar
g	gram
h	hours
HRP	horseradish peroxidase
HSL	homoserine-lactone
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
kb	kilobases
l	litre
LAM	lipoarabinomannan
LMM	lactate minimal medium
Mb	megabases
$\mu$ l	microlitre
$\mu$ M	micromolar
mg	milligram
min	minutes
ml	millilitre
M	molar
nm	nanometer
nM	nanomolar
NMR	nuclear magnetic resonance
OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	poly-ethylene glycol
pM	picomolar
s	seconds
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SOB	super optimal broth
TBS	Tris-buffered saline
TEMED	Tetramethylethylenediamine
V	volts
ZN	Ziehl-Neelson

## **Acknowledgements**

I am extremely grateful for all the assistance and support I have received in preparing this thesis. In particular I thank my principal supervisor, Professor Stephen Gillespie, for his encouragement, enthusiasm and patience, for lending me precious relevant literature, for a wealth of ideas, and for his personal support throughout my research and clinical training. Many thanks are also due my other supervisor, Professor Brian Henderson, who guided many aspects of the project and also provided numerous practical tips and assistance on the bench.

None of this work would have been possible without the extremely generous cooperation of Professor Mike Young and his group at the University of Wales Aberystwyth, namely Danielle Young, Galina Mukamolova, Obolbek Turapov, Chris Finan and Adriana Ravagnani. The first seven months of this project were spent in their laboratory. I am deeply indebted to them not only for allowing me bench space and resources, but also for many hours spent teaching a clinician enough about protein expression and Rpfs to allow me to continue with the project after I returned to London.

At University College London I worked in John Ward's laboratory in the Department of Biochemistry. Again I am grateful to him both for the bench space and for his expertise and willingness to spend time considering various aspects of my project. In the Department of Medical Microbiology at the Royal Free Hospital Tim McHugh has been a constant source of information, advice and practical assistance. At Birkbeck College I thank Andrei Okorokov and Claire Bagneris for their assistance with the large-scale protein production and at the Eastman Dental Centre Lindsay Sharp offered helpful advice about phage display. Many thanks are due to Professor Dhillon and the staff of the Histopathology

Department at the Royal Free Hospital for their input into the immunocytochemistry work. Finally, I thank the Medical Research Council, whose generous sponsorship allowed me the opportunity to undertake this project.

## **CHAPTER I**

### **INTRODUCTION**



## **1.1 Tuberculosis and latency – the clinical problem**

Tuberculosis (TB) is recognised as one of the major threats to human health. Approximately 8 million active cases of TB are diagnosed every year and there are over 3 million deaths. Furthermore it is estimated that about 1.9 billion people (about one third of the world's population) are latently infected. The term latency has been used by the Centers for Disease Control in the United States and the American Thoracic Society to refer to the clinical state where an individual is infected with *M. tuberculosis*, but has no clinical disease (American Thoracic Society 2000). These individuals are tuberculin reactive but have no active disease either on chest radiography or elsewhere. Most people exposed to *M. tuberculosis* mount a cell-mediated immune response which prevents the progress of the disease, and limits the organisms to the site of entry into the lung tissue and local lymph nodes (Ghon focus). However total elimination of the bacilli is very slow and unless the infection is treated, latent infection is believed to be the typical outcome. The organisms may then reactivate after many years and lead to active disease. On conservative estimates, 5-10% of latently infected immunocompetent individuals will develop active TB over their lifetimes: for those co-infected with HIV the figure rises to 5-10% per year of life. The adoption of a latent infection state makes sense for the organism: it is able to exist in a large proportion of the individuals in a group whilst ensuring its own transmission and continuation by causing overt disease in a few. *M. tuberculosis* is therefore well adapted to the human host, as pointed out by Bentrup & Russell (2001). It is this 'human reservoir' which poses the biggest problem of all to TB control.

The exact physiological and metabolic state of the tubercle bacillus in latent infection is not known for certain. It is postulated that the anaerobic and acidic conditions within closed lesions in the lung reduce bacterial metabolism. The organisms in this state are

known as 'persisters'. It is not known whether these cells are metabolising very slowly or whether metabolism is entirely halted. The term 'persistence' has been defined as 'a phenomenon whereby microorganisms which are drug-susceptible when tested outside the body are nevertheless capable of surviving within the body despite intensive therapy with the appropriate antimicrobial drug' (see the review by McKinney, 2000). There appear to be different subpopulations of persisters, as described by Hu & Coates (2003): they all survive the immune response ('immune persisters') but some are also refractory to drug treatment ('antibiotic persisters'). Immune persisters grow well on solid media but antibiotic persisters cannot (Hu *et al* 2000). The presence of persisters leads to the lengthy treatment required for TB: current therapy for active disease depends on the use of at least three drugs (rifampicin, isoniazid and pyrazinamide) for a minimum of six months. This strategy emerged as a result of a series of clinical trials conducted by the MRC and its international collaborators (see Fox *et al* 1999). Isoniazid, which inhibits mycobacterial cell wall synthesis, exerts its major activity against rapidly dividing cells. This is demonstrated clinically by a rapid fall in sputum viable counts and quickly renders the patient non-infectious (East African/British Medical Research Council 1973). Rifampicin and pyrazinamide have some activity versus persisters and are the best clinical sterilising agents (East African/British Medical Research Council 1973). Despite this, attempts to reduce duration of therapy below six months have failed. The long duration of therapy leads to problems in adherence and completing the correct course of treatment. This in turn may lead to the development of drug-resistant forms of disease resulting from serial monotherapy.

The presence of persisters is, therefore, fundamental to the problem of treating active TB satisfactorily. Eradicating latent infection effectively also requires treatments which can target them. Novel therapies which could do this might thus have the two-pronged

benefit of reducing length of treatment in active cases, and eradicating latent TB. One way of achieving this would be to activate the persistent organisms so that they become responsive to conventional chemotherapeutic agents. In order to do this a better understanding of the mechanisms of persistence, dormancy and reactivation is required.

## **1.2 Dormancy**

### *1.2.1 Definition*

The persisters which occur in tuberculous infection are often rather loosely referred to as 'dormant'. This is a term used to refer to microorganisms themselves, rather than to a clinical state. The most obvious example of a dormant bacterial cell is the bacterial spore, in which there is a reversible shutdown of metabolic activity. Different authors consider the term dormancy to describe a wide range of different states and there is no overall consensus on a definition. It has often been used to describe cells in different states that have simply reversibly lost the ability to proliferate (see Kell & Young 2000). Wayne chose the term 'non-replicating persistence' to describe the state of the bacilli in his model of dormancy in *M. tuberculosis* (Wayne & Hayes 1996). Wayne's model, and non-replicating persistence, are described in Chapter 5, section 5.2. Barer (2003) defines dormancy as 'a reversible state of low metabolic activity in a unit that maintains viability'. Given that the phenomenon of dormancy necessarily requires the concept of resuscitation, Kell and Young arrived at the following definition: 'In a state of low metabolic activity and unable to divide or form a colony on an agar plate without a preceding resuscitation phase' (Kell & Young 2000). This is the definition which has been used for the purpose of all work in this thesis.

### 1.2.2 Site and physiology of dormant organisms

Following inhalation, the bacilli enter the pulmonary alveolar macrophages and proliferate within them. The organism survives by disrupting phagosome acidification and preventing phagosome-lysosome fusion. Some bacilli may escape from infected macrophages and cause a transient bacteraemia with dissemination of viable bacilli to other parts of the body. Once infected, the chance of developing active disease is greatest immediately following infection, declining exponentially thereafter. Within the first year after infection, in immunocompetent individuals, the incidence of clinically significant disease is approximately 1.5% and the cumulative risk of disease during the first five years is 5-10% (Chiba & Kurihara 1979). However in the majority the infection is subsequently controlled by T cell and macrophage responses (see Parrish, Dick & Bishai 1998 for a review) and latent tuberculosis infection is the result.

The exact location of the dormant tubercle bacillus within the host is uncertain. Classically it was believed to reside in the macrophages in the tuberculous granuloma and healed pulmonary lesions (see Bishai 2000 for a brief review). These encapsulated and sometimes calcified lesions could protect the bacillus from the immune response and the anaerobic, acidic and lipid-rich environment might trigger a dormant state. However a significant proportion of cases of post-primary tuberculosis arise in organs other than the lungs, making it highly likely that the dormant bacilli reside in extra-pulmonary sites also, presumably having been seeded during the early bacteraemic phase of the primary infection described above. In 1927, Opie and Aronson prepared human tissue homogenates from unaffected portions of the lung from necropsy specimens and found that they contained viable *M. tuberculosis* infectious for guinea pigs – strongly suggesting that latent organisms exist in apparently normal tissue. A study by Hernandez-Pando and colleagues (Hernandez Pando *et al* 2000) found

evidence of *M. tuberculosis* DNA by *in-situ* and conventional PCR in 15 of 47 lung specimens taken from Mexican and Ethiopian patients who died of causes other than TB. The DNA was found to be present not only in macrophages, but also in type II pneumocytes, endothelial cells and fibroblasts. There was no histologic evidence of the presence of *M. tuberculosis*, either as histologic lesions or as visible acid-fast bacilli. There is other experimental evidence to suggest that acid-fast bacilli may not always be present in latent tuberculous infection. The Cornell model is an *in vivo* mouse model of dormancy in *M. tuberculosis* (McCune *et al* 1957a and 1957b). This model involves infecting mice with *M. tuberculosis* and then treating them with isoniazid and pyrazinamide for twelve weeks. On withdrawal of treatment, microscopy of tissue sections revealed that no acid-fast bacilli were visible, and cultivable tubercle bacilli could not be recovered from the tissues, even by passing tissue extract into other animals - a situation known as the 'sterile state'. However a few weeks later disease reappeared in the infected animals. More recently de Wit *et al* used PCR to detect the presence of DNA in the organs of sterile state Cornell model mice (de Wit *et al* 1995). Substantial amounts of DNA were present despite failure to obtain positive cultures. The presence of DNA could represent dead bacilli, free DNA or dormant forms. However studies using RT-PCR showed that mRNA transcripts and rRNA are present in sterile state mice in the Cornell model, implying that these non-culturable bacilli are transcriptionally active (Hu *et al* 2000, Pai *et al* 2000).

Thus the absence of visible acid-fast bacilli on microscopy certainly does not seem to preclude their presence in the tissues in some form. This has prompted some to consider the possibility of a non-acid-fast form of *M. tuberculosis* (eg 'ghost cells' -Canetti 1955, Seiler 2003). *M. tuberculosis* does possess genes homologous to sporulation regulatory genes of actinomycetes. This may explain the ability of *M. tuberculosis* to survive for

such long periods in latent infection due to the putative induction of a 'spore-like state'. Others argue that where acid-fast bacilli are not visible in infected tissues this may simply be down to the lack of sensitivity of microscopy, and that just because *M. tuberculosis* possesses sporulation regulatory gene homologues does not mean they are activated (Orme 2001). Isoniazid significantly reduces the risk of reactivation in latent tuberculosis, but it acts only on actively metabolising bacilli (as a cell wall agent) - this observation does not support the presence of a totally metabolically inert spore-like state. A further discussion of the location and activity of dormant tubercle bacilli is found in the introduction to Chapter VII.

### 1.2.3 *Molecular changes in dormancy*

The molecular changes leading to dormancy are also beginning to be unravelled. Transcriptional regulators such as sigma factors are likely to have an important role in bacterial adaptive responses. *SigF* is an alternate sigma factor found in *M. tuberculosis* and is similar to sporulation sigma factors in *Bacillus subtilis*. It has been found to be induced in *M. tuberculosis* by a number of stresses including anaerobic conditions, oxidative stress, antibiotic stress and nutrient depletion (Michele, Ko & Bishai 1999). It is therefore believed that *sigF* may control genes which are important for mycobacterial persistence within the host, including during chemotherapy. The fact that expression can be induced both by anaerobiosis and antibiotic treatment lends support to the idea that dormancy and persistence may be very closely related. It is not certain whether the expression of *SigF* indicates a stress response or the presence of a sporulation cascade. Electron microscopy of tubercle bacilli exposed to hypoxia shows thickening of the cell wall and analysis of the cultures by SDS-PAGE identified the presence of a 16kDa alpha-crystallin (Acr) protein (Cunningham & Spreadbury 1998). This protein was strongly associated with the cell envelope, reminiscent of a sporulation-like response,

and the authors suggested that it might play a role in stabilizing cell structures during long-term survival in granulomas. The *acr* gene is strongly induced by hypoxic conditions and is required for growth in macrophages (Yuan *et al* 1998).

Although *M. tuberculosis* is usually considered a strict aerobe, it can survive in anaerobic conditions (Wayne 1976) and to this end it is able to use fatty acids as a carbon source: the glyoxylate shunt is upregulated (Wayne & Lin 1982). This shunt requires the enzyme isocitrate lyase (ICL), and it has been found that persistence of *M. tuberculosis* in mice is facilitated by ICL, with disruption of the *icl* gene causing attenuation of bacterial persistence and virulence in the mice (McKinney *et al* 2000). Under anaerobic conditions the organism also switches to nitrate respiration, with a marked increase in nitrite production accompanying anaerobic persistence (Wayne & Hayes 1998).

A detailed review of the molecular changes identified in dormancy is beyond the scope of this introduction, which will concentrate mainly on Resuscitation Promoting Factors, but changes such as those briefly described all suggest that the metabolism and surface structure of *M. tuberculosis* are profoundly influenced by the environmental conditions.

### **1.3 Resuscitation promoting factors.**

Some aspects of dormancy have been briefly reviewed in this introduction. It will be apparent that much remains to be elucidated about the nature of dormancy, its causes and also about what causes organisms to come out of the dormant state. One very intriguing area pertaining to the latter question is that of the resuscitation-promoting factors (Rpfs). Since the Rpfs are the subject of this study, the literature surrounding them will now be reviewed in some detail.

### 1.3.1 Discovery of *Rpf*

*Micrococcus luteus* is a high G+C Gram-positive organism that may be isolated from mammalian skin, soil and water. In 1994 Kaprelyants *et al* published a study in which it was found that dormant (ie after a prolonged stationary-phase) cells of *M. luteus* in liquid medium could be stimulated into growth by adding supernatant from stationary phase cultures of the same organism. This was evidenced using the most probable number method, which showed an increase in apparently viable cells of 1000–100 000-fold in cultures with the added supernatant, compared to those without it. It was concluded that the spent culture supernatant contained a factor which stimulated the growth of dormant cells (Kaprelyants *et al* 1994, Mukamolova *et al* 1995).

It was established that the active component of the supernatant was heat-labile, nondialyzable, and trypsin-sensitive (Mukamolova *et al* 1998). It was purified to homogeneity by ion-exchange chromatography. The peak resuscitation activity corresponded to a protein with a molecular mass of 16-17kDa, named resuscitation-promoting factor (Rpf). It was found to be active in the picomolar range and increased the number of culturable *M. luteus* cells from dormant populations by three orders of magnitude. In view of the low concentrations necessary for activity a nutritional role for Rpf was discounted. Primers designed from microsequence data were used to amplify a 147bp fragment from *M. luteus* DNA. When this fragment was cloned and used as a hybridization probe it detected a 1.4kbp fragment, corresponding to a gene denoted *rpf* (Mukamolova *et al* 1998).

Data from bacterial genome sequencing projects have now revealed the existence of over 40 examples of *rpf*-like genes in the high G+C cohort of Gram-positive bacteria, including streptomycetes, corynebacteria and mycobacteria (Mukamolova *et al* 1998,



Kell & Young 2000, Ravagnani, Finan & Young 2005). Most organisms contain several representatives (unlike *M. luteus* which only has one). Recombinant *M. luteus* Rpf has a wide spectrum of activity - it stimulates the growth of several other high G+C Gram positive organisms, including *Mycobacterium tuberculosis*, *Mycobacterium smegmatis* and *Mycobacterium avium* (Mukamolova *et al* 1998). An Rpf domain has even been identified in a large mycobacteriophage (Pedulla *et al* 2003).

### 1.3.2 Structure of Rpf and possible modes of action

The *rpf*-gene encodes a product of total length 223 amino acids. Bioinformatic analysis suggests it has two functional domains. At the N-terminus, the first 41 amino-acid residues are predicted to be a secretory signal peptide, which is removed during protein processing to give the 182 amino-acid mature protein, sequestered on or in the bacterial cell wall. At the C-terminus, a LysM module is found, which probably promotes its association with the peptidoglycan of the cell envelope (Bateman & Bycroft 2000). The Rpf domain also contains two highly conserved cysteine residues believed to lie in close proximity and to form a disulphide bridge (Ravagnani, Finan & Young 2005). All the *rpf*-like genes share extensive sequence similarity over a 70aa segment, known as the 'Rpf domain' (Figure 1.1). A truncated form of the protein consisting only of the 'Rpf domain' is fully active and the C-terminus LysM module is not required for biological activity (Mukamolova *et al* 2002a).

**Figure 1.1** Multiple sequence alignment of the conserved region of *M. luteus* Rpf and the five Rpf orthologues of *M. tuberculosis*. Conserved residues are in red and residues in green represent those absolutely conserved in all known Rpf orthologues.

Rv0867c (RpfA)	EWDQVAR CESGGNWS I NTGNGYLGGGLQFTQS TWASH	91
Rv1009 (RpfB)	I WDA I AGCEAGGNWAI NTGNGYYGGVQFDQGTWEAN	332
Rv1884c (RpfC)	NWDAVAQ CESGGNWAANTGNGKYGGGLQFKPA TWAAF	117
Rv2389 (RpfD)	DWDA I AQCESGGNWAANTGNGLYGGGLQ I SQATWDSN	98
Rv2450c (RpfE)	NWDA I AQCESGGNWS I NTGNGYYGGL RFTAGTWRAN	145
Rpf	TWD R LAECESNGTWD I NTGNG FYGGVQFTLS SWQAV	88
Rv0867c (RpfA)	GGGEFAPSAQLASREEQ I AVGERVLATQGRGAWPVCG	140
Rv1009 (RpfB)	GGLRYAPRADLATREEQ I AVAEVTRLRQGWGAWPVCA	362
Rv1884c (RpfC)	GGVG --- NPAAASREEQ I AVANRVLAEQGLDAWPTCG	166
Rv2389 (RpfD)	GGVG --- SPAAASPQQQ I EVADNIMKTQG PGAWPKCS	147
Rv2450c (RpfE)	GGSG --- SAAN ASREEQ I RVAENVLRSQG I RAWPVCG	172
Rpf	GGEG --- YPHQASKAEQI KRAE I LQDLQGWGAWPLCS	137

Rpf was considered in 1998 to be the first autocrine or paracrine factor identified which stimulated the resuscitation of dormant bacteria (Mukamolova *et al* 1998); it was considered to have the properties of a cytokine because it also stimulated the growth of viable cells and was thought likely to have a role in the normal control of cell multiplication. The high potency of Rpfs, and the fact that there is an optimum concentration for activity, above which they are inactive (a phenomenon seen in studies of ligand-receptor interactions) seemed to support the suggestion that Rpfs bind to specific protein receptors. Over the last twelve months however evidence has been mounting that the Rpfs may have a different mode of action, namely, as muralytic enzymes. Alignments show a 30% homology between the Rpf domain and the c-type lysozyme family (Cohen-Gonsaud *et al* 2004). The conserved residues are glutamate at position 35, glycine at positions 46 and 104, and tryptophan at positions 64 and 108. All

deletions appear to be sited in relatively unimportant parts of loops, and although the cysteine residues are not conserved in lysozyme, the disulphide bridges are. Hence, the Rpf domain has a lysozyme-like fold. The molecular structure of the common Rpf core, or Rpf domain, as determined by NMR spectroscopy, was recently published. (Cohen-Gonsaud *et al* 2005). This confirmed the resemblance of Rpf to glycoside hydrolases and the presence in Rpf of a cleft equivalent to the oligo-saccharide binding site cleft of c-type lysozymes. Using NMR, binding of the Rpf domain to trimers of the polysaccharide NAG was demonstrated in the form of chemical shift changes for residues of the putative binding groove. The experiments suggested that the same binding mechanism is present in c-type lysozymes as in Rpf. The need for an active site glutamate was confirmed by producing a mutant Rpf which lacked this residue: the resuscitation activity of the mutant protein was completely suppressed. Hence, although the overall sequence conservation between Rpf and c-type lysozymes is low, the need for an active site glutamate and a binding pocket that undergoes rearrangement in the presence of a polysaccharide strongly suggest that the function of Rpfs may be to cleave oligosaccharide in the cell wall. Another publication from Mike Young's group supports this view (Ravagnani, Finan & Young 2005). In this work, the Rpf proteins were classified into several subfamilies by *in silico* analysis of the accessory domains. The RpfB subfamily has very similar domain structures and genomic contexts to a group of proteins of unknown function in the low G+C Gram positive bacteria (which do not contain Rpfs). In these proteins, the Rpf domain is replaced by another completely different domain designated Sps (stationary phase survival). Forty-six Sps domains have been identified. Organisms containing *sps*-like genes include *Bacillus anthracis*, *Bacillus cereus*, *Listeria monocytogenes*, *Enterococcus faecalis*, *Clostridium botulinum*, *Clostridium tetani* and *Clostridium perfringens*. The Sps domain was found to be associated with other domains known to be present in muralytic enzymes (eg

SH3b and LysM). The sequence similarity between the C-terminal region of the Sps domain and that of the Gram-negative lytic transglycosylase MltA also supports this association. Mike Young's group have subsequently shown that *M. luteus* Rpf has murein hydrolase activity in zymograms (Mukamolova et al 2006). This has led to their suggestion that the walls of dormant organisms may contain inert peptidoglycan which could be likened to a 'cocoon', requiring the muralytic enzymes to cleave certain residues so that growth and cell wall expansion can resume. It is still possible that cell-cell signalling might be involved, perhaps mediated by a small molecule released as a result of the cleavage of peptidoglycan by Rpf.

### 1.3.3 Activity of Rpf

Once the gene encoding the Rpf had been identified, it was cloned and expressed in *E. coli* (Mukamolova *et al* 1998). The recombinant protein was found to have the following activity:

- i. it promoted the resuscitation of dormant cells of *M. luteus*, in a similar manner to the culture supernatant used in the first experiments. Activity is greatest when the proteins are used in picomolar concentrations; activity is lost at both higher and lower concentrations.
- ii. it shortened the apparent lag phase (ie as estimated by optical density of the culture medium) of *M. luteus* in batch culture.
- iii. it enabled small inocula of washed cells to grow in minimal media, that could not grow in the absence of Rpf.

Rpf was also able to stimulate the growth of washed *Mycobacterium smegmatis* cells in Sauton's medium (a relatively minimal medium). The cells could not grow in the absence of Rpf. Similarly, Rpf stimulates the growth of *M. tuberculosis*, *Mycobacterium*

*avium*, *Mycobacterium bovis* BCG and *Mycobacterium kansasii* in Sauton's medium, as evidenced microscopically after Ziehl-Neelson staining of 50 µl culture after 14 days.

Rpf has been described as an essential secreted growth factor (Mukamolova *et al* 2002a). As well as being essential for growth of washed cells of *M. luteus* inoculated at low density into a minimal medium, incorporation of anti-Rpf antibodies into the culture medium at the time of inoculation also prevented bacterial growth. Addition of recombinant Rpf as well as anti-Rpf antibodies to the medium abolished this inhibition. Furthermore, attempts to disrupt the *rpf* gene failed except in the presence of a second functional plasmid-located copy of the gene. An *rpf* null mutant could not be constructed, even when recombinant Rpf was added to the growth medium. On the other hand, Hartmann (Hartmann *et al* 2004) studied *Corynebacterium glutamicum* which contains two *rpf*-gene homologues. In this organism, strains with deletions of both these genes could be constructed, indicating that the two genes were not collectively essential for *C. glutamicum*. However the double mutant displayed slower growth and a prolonged lag phase after transfer of long-stored cells into fresh medium.

#### 1.3.4 Onset of transcription and appearance of Rpf in cultures

Isolation of mRNA at various times in the *M. luteus* bacterial growth cycle revealed that transcripts of Rpf became detectable by RT-PCR within 30 min of inoculation of the medium with *M. luteus* cells. Within 1 hour of inoculation, transcripts were abundant, and they remained so throughout the lag phase. From early mid-log phase onwards the transcript abundance began to decline, until they were undetectable in late stationary phase. To measure the accumulation of Rpf itself in the medium, ELISAs were used, which revealed that Rpf was not detectable until about 5 hours after inoculation, by which time almost two doublings had occurred. Rpf accumulated in the culture medium

during early mid-exponential phase and declined afterwards, reaching zero at 50 hours post-inoculation. Hence there was a delay between the onset of *rpf* transcription and the appearance of Rpf in the culture supernatant. The probable explanation for this was given by examining the bacteria by confocal microscopy after labelling with anti-Rpf antibodies. Rpf was present on the bacterial cell surface in mid-exponential phase: in samples taken much earlier during the growth cycle Rpf was present on a higher number of cells (~30%). Many of the protein molecules may have remained cell-associated, to be released later into the culture medium. This event might represent a signal, anticipating onset of a period of less active growth in preparation for stationary phase, or, the action of Rpf might be mediated whilst still attached to the cell surface.

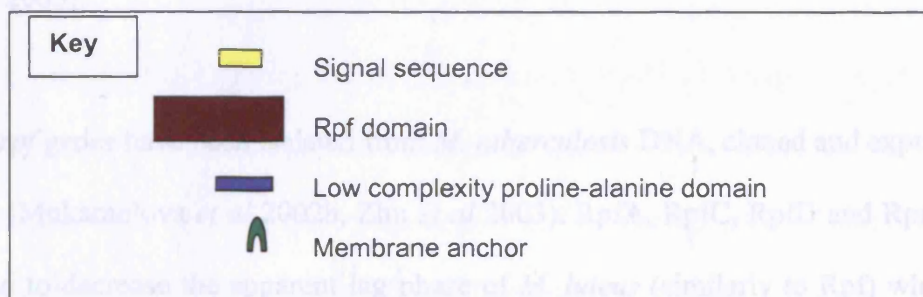
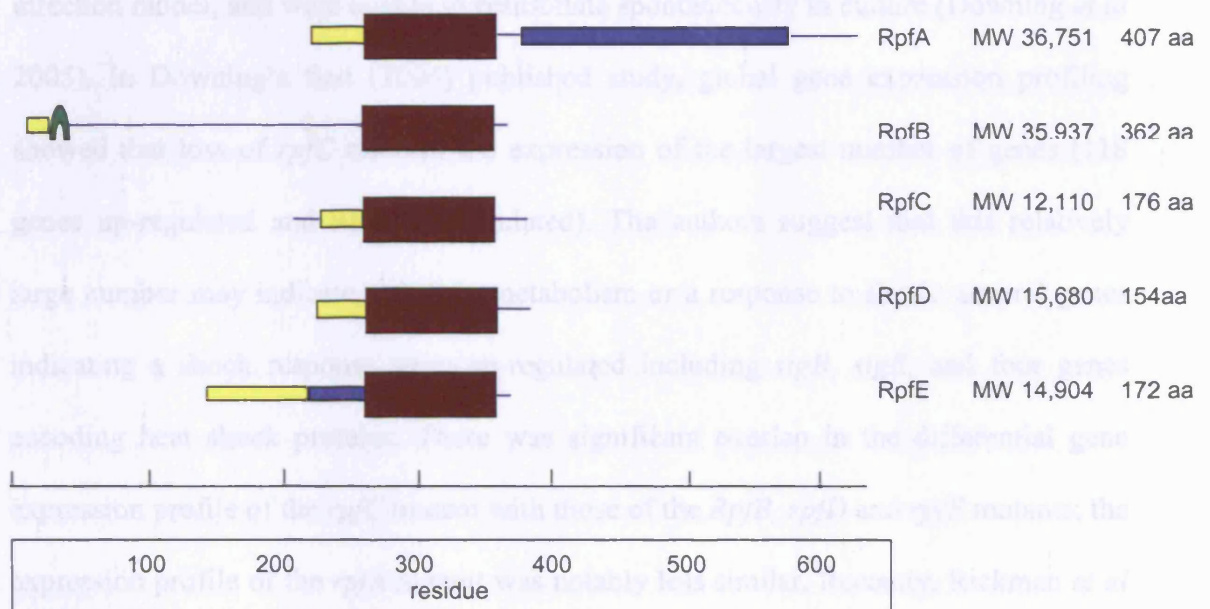
#### 1.3.5 *Rpfs in M. tuberculosis*

Homology searches revealed that *M. tuberculosis* contains five *rpf* gene orthologues. All show extensive sequence similarity to the N-terminal (signalling) domain of Rpf (See Figure 1.1). Two of the proteins, RpfA (Rv0867c, 407 aa) and RpfD (Rv2389c, 154aa) are predicted to be secreted. RpfA is a comparatively large protein in which the Rpf-like sequence is followed by a long series of proline and alanine-rich repeats at residues 146-320. RpfB (Rv1009, 362aa) seems to have a membrane anchor at the N-terminus and its Rpf-like domain at the C-terminus. It is thought to be anchored to the outer surface of the cell membrane by an N-terminal prokaryotic membrane lipoprotein lipid attachment site. The status of the products of the other two genes, RpfC (Rv1884c, 176aa) and RpfE (Rv2450c, 172aa), is less clear though they are probably also membrane-associated or secreted (Mukamolova *et al* 2002b). Like Rpf from *M.luteus*, all five probably have extra-cytoplasmic functions. Figure 1.2 shows the postulated functional regions of the *rpf*-like gene products from *M. tuberculosis*, as a schematic



comparison. The gene and amino acid sequences of each protein, and their molar extinction co-efficients, are shown in Appendix i.

**Figure 1.2** Diagrammatic representation of all five *M. tuberculosis* Rpf s, showing predicted molecular weights and number of amino acid residues. Adapted from Mukamolova *et al* 2002b and Ravagnani, Finan & Young 2005.



The five genes are scattered around the *M. tuberculosis* genome. This protein family may show functional redundancy, or, the different signalling molecules may be produced under different conditions or for different specific purposes. All five knock-out mutants for each one of RpfA-E of *M. tuberculosis* are viable (Downing *et al* 2004, Tufariello *et al* 2004): the cells grew as well as wild type in aerated liquid medium, although a slightly lower final cell density was reached and colonies were slightly smaller on agar. However, mutants of *M. tuberculosis* lacking three of the *rpf* genes (*rpfA*, *rpfB* and *rpfC* or *rpfA*, *rpfC* and *rpfD*) were significantly attenuated in a mouse infection model, and were unable to resuscitate spontaneously in culture (Downing *et al* 2005). In Downing's first (2004) published study, global gene expression profiling showed that loss of *rpfC* affected the expression of the largest number of genes (118 genes up-regulated and 91 down-regulated). The authors suggest that this relatively large number may indicate a shift in metabolism or a response to shock: several genes indicating a shock response were up-regulated including *sigB*, *sigE*, and four genes encoding heat shock proteins. There was significant overlap in the differential gene expression profile of the *rpfC* mutant with those of the *RpfB*, *rpfD* and *rpfE* mutants; the expression profile of the *rpfA* mutant was notably less similar. Recently, Rickman *et al* have shown that RpfA expression is regulated by a transcription factor belonging to the cAMP receptor protein family. Deletion of Rv3676, which encodes this transcription factor, causes growth defects both in laboratory culture and in a mouse model (Rickman *et al* 2005).

The *rpf* genes have been isolated from *M. tuberculosis* DNA, cloned and expressed in *E. coli* (Mukamolova *et al* 2002b, Zhu *et al* 2003). RpfA, RpfC, RpfD and RpfE were all found to decrease the apparent lag phase of *M. luteus* (similarly to Rpf) when used in subpicomolar concentrations. RpfB was not tested in this context. RpfA and RpfC were



the most potent: RpfA showed activity at fM concentrations (equivalent to just a few molecules per cell), whilst RpfC caused the greatest reduction in apparent lag phase (from 216h to 80h). As with Rpf, there appears to be an optimum concentration range for activity, above and below which there is reduced activity or none at all. The proteins appear to lose biological activity during storage.

The five *M. tuberculosis* Rpfs were also tested on cells of *M. bovis* BCG, which contains five *rpf* gene homologues very similar to those of *M. tuberculosis*. When actively growing cells were used, the Rpf proteins did not stimulate bacterial growth. However, in common with *M. tuberculosis*, cells of *M. bovis* BCG lose culturability during extended stationary phase (Sun *et al* 1999, Lim *et al* 1999, Shleeve *et al* 2002); see Chapter 5. Use of such late stationary phase cells showed that all five proteins, used in pM concentrations, enabled dormant cells to grow that could not grow in their absence (Mukamolova *et al* 2002b). In the experiments on *M. luteus*, as stated previously, it was possible to produce a state in which cells could not grow without the addition of Rpf, by washing the cells and then inoculating them at low density in minimal medium. Washing presumably removes secreted Rpf associated with the cell envelope. However it was not possible to reproduce this effect with *M. tuberculosis* and *M. bovis*, supporting the belief that one or more of the Rpfs produced by these organisms is membrane bound (specifically, RpfB).

To investigate the expression of the five genes, mRNA corresponding to each was detected by RT-PCR using RNA extracted from *in vitro* growing cultures of *M. tuberculosis* and *M. bovis* BCG (Mukamolova *et al* 2002b). All five were expressed in both organisms under these conditions. RpfC appeared much more highly expressed than the others (unpublished data). However, mRNA corresponding to the genes could

not be detected in cultures of *M. bovis* BCG during stationary phase, or from cells which had been starved for 5 months. These Rpf genes are thus expressed in actively growing cells, but not in non-growing cells (or they are below the limit of detection). Rpf proteins were also detected in the supernatants of actively growing cultures of *M. tuberculosis* and *M. bovis* BCG, using anti-Rpf antibodies. As with the *M. luteus* experiments, use of confocal microscopy of *M. bovis* BCG incubated with anti-Rpf antibodies and a secondary FITC-labelled antibody showed that Rpf-like proteins could be detected on the bacterial cell surface. About 30% of the cells in early logarithmic phase cultures showed fluorescence. None was seen in stationary phase cultures. Adding anti-Rpf antibodies to the culture medium inhibited the growth of both *M. bovis* BCG and *M. tuberculosis*; this effect could be abolished by also adding Rpf.

#### 1.3.6 Effect of Rpf expression *in vivo*

As stated earlier, recombinant proteins appear unstable with storage and the number of biologically active molecules in each batch of protein may vary widely. Because of this attempts have been made to examine the effect of Rpf expression *in vivo* (Mukamolova *et al* 2002b, Shleeve *et al* 2004). In Mukamolova's experiment, a plasmid expressing *rpf* under the control of a *M. smegmatis* promoter was introduced into *M. smegmatis*. When these bacteria were subjected to inoculation at low density into a nutrient-poor medium (Sauton's) the apparent lag phase was reduced in the Rpf-expressing strain compared to a control. The effect was not visible at a high inoculation density. Shleeve found that dormant cells of *M. smegmatis* harbouring plasmids expressing *rpf* resuscitated spontaneously in Sauton's medium after an apparent lag of 3 days. Control cells with the same plasmids but lacking *rpf* failed to resuscitate spontaneously, as had been observed with the wild-type.

### 1.3.7 Clinical relevance of Rpf

The study of the *M. tuberculosis* Rpfs may have great clinical relevance. It has been demonstrated that *M. tuberculosis* cells isolated from murine macrophages, which could not be cultured under normal conditions, could be resuscitated using Rpf in liquid medium (Biketov *et al* 2000). This has implications for the laboratory diagnosis of tuberculosis, where cultures may be falsely negative or very slow-growing because of failure of growth of organisms which have become dormant or metabolically inactive. Addition of Rpf or Rpf analogues to culture media might lead to more sensitive and rapid culture of slow-growing mycobacteria – for a preliminary study see Freeman *et al* 2002. From a therapeutic point of view, cells in a low state of metabolic activity are less susceptible to drug treatment, so to resuscitate them might be to render them easier to treat with conventional drug treatments. Rpfs might be used as an adjunct to conventional treatment and thus result in a shorter course of chemotherapy. Conversely, where there is a high risk of reactivation of tuberculosis, for example in HIV infection, suppression of Rpf expression or activity might be able to prevent the disease becoming active again. Finally, *Micrococcus luteus* Rpf and RpfE have been used in mice to produce an immune response against infection with *M. tuberculosis* (Yermeev *et al* 2003). IgG1 and IgG2 responses were elicited, as well as T-cell proliferation and the production of gamma interferon, IL-10, and IL-12. The levels of protection obtained, as evidenced by survival times and mycobacterial multiplication in organs, was comparable to that obtained using short-term culture filtrate components such as Ag85B and ESAT-6, or by using a low dose ( $10^3$  cfu) of BCG: although lower than the level achieved by vaccination with  $10^4$  –  $10^6$  cfu of BCG.

## 1.4 Aims of the project

The original aim of this project was to investigate the effects of the mycobacterial Rpf's on the antimicrobial susceptibility of dormant cultures of *M. tuberculosis*. It was postulated that the resuscitation of dormant cultures of the organism, using Rpf's, would simultaneously increase their susceptibility to anti-tuberculosis drugs, which is known to be reduced in the dormant state. However, difficulties with the stability and solubility of the Rpf's, and the measurement of their activity, precluded this. Hence the first part of the project, described in chapters 3-5, concentrated on resolving these problems by improving the expression and solubility of recombinant Rpf's, and developing assays for measuring their activity, to pave the way for future work. Chapter 6 is an account of the attempt to identify a protein receptor molecule for Rpf's in *M. tuberculosis*, using the technique of phage display. Chapter 7 describes the location of Rpf's in *M. tuberculosis*-infected human tissue sections, as identified by immunocytochemistry.

## **CHAPTER II**

### **MATERIALS AND METHODS**

## **2 Materials and Methods**

All chemicals were obtained from Sigma and BDH, unless stated otherwise in the text. Media and solutions are listed in Appendix *ii*.

### **2.1 Cloning experiments**

#### *2.1.1 DNA extractions from Escherichia coli*

Strains for DNA extraction were taken from the  $-80^{\circ}\text{C}$  freezer and cultured overnight in Luria-Bertani broth (LB) or Nutrient Broth 2 (NB2) containing the appropriate concentration of antibiotic, in 2 x 2 ml glass test tubes or in sterile plastic universals on a shaking incubator at  $37^{\circ}\text{C}$ . The following day DNA was extracted using the Qiagen QIA Prep Spin Miniprep kit, following the manufacturer's instructions.

#### *2.1.2 DNA extractions from M. tuberculosis*

This was carried out in the containment level 3 laboratory. A loopful of cells of H37Rv from a Lowenstein-Jensen slope was transferred into a sterile 1 ml microcentrifuge tube containing 400  $\mu\text{l}$  of TE buffer. This was heated at  $80^{\circ}\text{C}$  for 20 min to kill the cells, then cooled to room temperature. 50  $\mu\text{l}$  of 10mg/ml lysozyme were added, the suspension vortexed briefly and incubated for 1 hour at  $37^{\circ}\text{C}$ . 75  $\mu\text{l}$  10% SDS/proteinase K solution was added, the tube inverted, and incubated for 10 min at  $65^{\circ}\text{C}$ . 100  $\mu\text{l}$  of 5 M NaCl was added, then 100  $\mu\text{l}$  of CTAB/NaCl solution (prewarmed at  $65^{\circ}\text{C}$  for 5 min). The liquid was vortexed until it became white, then incubated for 10 min at  $65^{\circ}\text{C}$ . 750  $\mu\text{l}$  chloroform/isoamyl alcohol was added, the tube mixed by inversion and centrifuged at room temperature for 5 min at 12,000 g.

The rest of the procedure was carried out outside the containment level 3 laboratory. All the aqueous supernatant was transferred, in 180  $\mu\text{l}$  aliquots, to a fresh 1 ml

microcentrifuge tube, using a pipette. Next 450 µl ice-cold isopropanol was added to precipitate the nucleic acids and the tube rolled carefully by hand. It was then placed at -20°C for 30 min, before microfuging for 15 min at 10,000 g. Most of the supernatant was removed, and the pellet washed with 1 ml ice cold 70% ethanol, and inverted gently. This was microfuged at 10,000 g for 5 min, and the supernatant discarded carefully. The pellet was allowed to air-dry at room temperature and resuspended in 50 µl TE buffer, overnight at 4°C.

### 2.1.3 Measurement of DNA concentrations

This was done with the PicoGreen dsDNA Quantitation kit (Molecular Probes), as per the manufacturer's instructions.

### 2.1.4 PCR for genes *rpfD* and *Rv0328*

Primers were designed for the *rpf-D* gene that did not include the N-terminal signal sequence. The primers were designed to have an *NdeI* site at the N-terminal/5' end (primer sequence CGC ATA TGG ACG ACA TCG ATT GGG ACG CC, restriction site underlined) and an *XhoI* site at the C-terminal/3' end (primer sequence CTC GAG ATC GTC CCT GCT CCC CGA ACA A). The sequence for amplification had a length of 315 bp. The reaction was carried out with 5 µl *M. tuberculosis* DNA (concentration 50 µg/ml) both in the presence and absence of Q solution (Qiagen) in a total reaction volume of 50 µl. 1 µl Taq polymerase was added after heating to 95°C for a 'hot start'.

Primers for *Rv0328* were designed with an *NdeI* site at the 5' end (primer sequence ATA TCA TAT GCA ACA GCA ACG CAC A) and a *HindIII* site at the 3' end (primer sequence ATA TAA GCT TCA CCG CAA CTG CGA CGT). The reaction was carried

out as for the *rpfD* gene PCR but without Q solution and without using 'hot start'. The PCR protocol was:

5 min @ 95°C

30 cycles: Denaturation: 95 °C for 30 s

Annealing: 60 °C for 90 s

Extension: 72 °C for 120 s

Final extension: 72°C for 10 min

#### *2.1.5 TOPO 2.1 ® cloning reaction*

This was done according to the manufacturer's instructions (Invitrogen).

#### *2.1.6 Restriction endonuclease digestions*

Enzymes for digestions, and their respective buffers, were obtained from New England Biolabs (NEB). Buffers were used at a concentration of 10% of the total reaction volume (v/v), which was usually 20 µl. Digestions were carried out in 1.5 ml microcentrifuge tubes. Where *Bam*HI was used bovine serum albumin (NEB) was also added to the reaction mix (0.2 µl) Each digestion reaction was incubated at 37°C for 90-120 min, then examined by agarose gel electrophoresis.

#### *2.1.7 Agarose gel electrophoresis*

This followed the standard method for a 1% agarose gel in TAE buffer. 1 g agarose was dissolved in 100 ml TAE buffer by heating in a microwave oven. The agarose was cooled to about 50°C, then ethidium bromide (0.5 µg/ml) was added. The mixture was then poured into a taped gel tray containing the appropriate comb. After cooling, the solid gel in the tray was transferred to a horizontal gel tank filled with TAE. DNA samples containing 10% (v/v) loading buffer were loaded into the wells left after



removal of the comb and the tank was connected to a power supply. This was run constantly at 10-15 V/cm, and the DNA was left to migrate through the gel towards the positive pole. The ethidium bromide-stained DNA was visualised under ultraviolet light (302 nm). Where small DNA fragments were to be visualised a 2% gel was used.

#### *2.1.8 Extraction of DNA from agarose gel*

For this procedure, the desired band of DNA was excised from the agarose gel with a clean scalpel under ultraviolet light, then extracted from the gel using the QIAEX II Agarose Gel Extraction Kit (Qiagen), following the manufacturer's instructions.

#### *2.1.9 Treatment with calf-intestinal alkaline phosphatase*

On occasion digested DNA was treated with calf-intestinal alkaline phosphatase (CIP) prior to ligation. 1 µl calf-intestinal CIP (NEB), 1 µl 10 x buffer (NEB) and 8 µl distilled water were added to 30 µl of digested DNA and incubated at 37°C for 1 h. The enzyme was removed from the reaction mixture by passing through a QIA quick column (Qiagen), before setting up the ligation itself as above.

#### *2.1.10 Ligation reactions*

T4 DNA ligase (NEB) was used with 10 x T4 DNA ligase buffer (NEB). 1 µl of T4 DNA ligase and 1 µl buffer were used with the required DNA in a 10 µl total volume reaction. The reaction was performed in a 1.5 ml microcentrifuge tube and incubated overnight on the bench.

Alternatively in some cases Ready-To-Go™ T4 DNA ligase (Amersham Biosciences) was used. 20 µl of DNA solution was added to the tube, which was incubated at room temperature for 3-5 min, then gently mixed by pipetting. The tube was incubated at

room temperature for a further 30-45 min, then the ligase was inactivated by heating to 70°C for 10 min.

#### *2.1.11 Preparation of competent E. coli host cells*

5 ml of nutrient broth was inoculated with the *E. coli* host cells and grown on a shaker-incubator at 37°C for 3-4 hours until the OD<sub>600</sub> was about 0.6. 1.5 ml of the culture was pelleted in a microcentrifuge tube (14 000 rpm for 5 min) and then resuspended by pipetting in 0.75 ml of an ice-cold solution of 0.075 M CaCl<sub>2</sub>/15% glycerol. The cells were pelleted again and gently resuspended in 0.15 ml of the same solution. They were then used immediately, or stored at -70°C.

#### *2.1.12 Transformation of competent cells*

200 µl of the competent cell suspension was mixed with 1 µl DNA, using a cold pipette tip. This mixture was left on ice for 10 min and then heat-shocked at 42°C for 90 s. SOB (800 µl) was added and the mixture was shaken at 37°C for 90 min. 0.1 – 0.2 ml aliquots of the transformed cells were plated onto selective antibiotic plates with a plate spreader and incubated at 37°C overnight.

#### *2.1.13 Blue-white selection of XL2-Blue strain (Novagen)*

This method was used in the selection of XL2-Blue cells transformed with pET-43.1b containing the Rpf<sub>smeg</sub> gene insert and of DH5α cells transformed with the TOPO 2.1® vector containing the *rpfD* gene insert or the Rv0328 gene insert.

50 µl, 100 µl and 200 µl aliquots of the transformed cells were plated onto LB plates containing ampicillin (50 µg/ml), 40 µM 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG).

The plates were incubated overnight at 37°C. White colonies were picked off and subcultured onto LB plates containing 50 µg ampicillin/ml and re-incubated for 6 h. These subcultured colonies were then incubated in broth overnight on a shaker. Next day, a subculture from each broth was made on LB/Ap50, streaking to obtain single colonies.

#### *2.1.14 Sequencing of cloned gene inserts*

This was done by thermal cycle sequencing, using the ABI 310 system, a capillary electrophoresis-based sequencer using laser fluorescence detection of Big Dye terminators. Sequencing was carried out following the manufacturer's instructions.

#### *2.1.15 Storage of expression hosts and clones*

All hosts and clones were frozen in 20% glycerol and stored at -70°C.

## **2.2 Protein expression**

### *2.2.1 IPTG induction*

10 ml Nutrient Broth E (NBE, Lab M) or NB2 (Oxoid) containing 50 µg ampicillin/ml in a 100 ml conical flask, was inoculated with the desired bacterial strain from the -70°C freezer, and cultured on a shaker-incubator at 37°C overnight. The following morning, 1.5 ml of the culture was removed and centrifuged for 1 min at 13,000 rpm in a microcentrifuge tube. The supernatant was discarded, to remove any β-lactamase that had accumulated in the overnight cultures.

500 ml conical flasks containing 150 ml NBE/NB2 and 50 µg ampicillin/ml were prepared. The cell pellet after centrifugation was resuspended in 1 ml of this broth and then inoculated into the 500 ml flask. The bacteria were cultured on a shaker-incubator

at 37°C until the culture had reached an OD<sub>600nm</sub> of between 0.7 and 0.9 (about 4 h). At this point, 1 mM IPTG was added and the incubation continued for a further 3 h. In the case of RpfC and RpfD, which have low levels of expression in these hosts and vectors, 300-ml cultures were used rather than 150-ml cultures. They were also incubated for 6 h after the addition of IPTG, whilst the other Rpfs only required incubation for 3 h. In some other experiments, the concentration of IPTG and the length and temperature of induction were varied to improve solubility and where this was done it is indicated in the text.

### *2.2.2 Harvesting the induced cells*

At the end of the induction, the cultures were centrifuged for 10 min at 6000 rpm. The cells were then resuspended in 50 ml 20mM Tris-HCl pH 8 and centrifuged again. The supernatant was discarded and the cell pellets stored at -20°C.

## **2.3 Protein extraction and purification**

### *2.3.1 Sonication*

When using urea, the frozen induced cell pellet was resuspended in 5 ml binding buffer containing 8 M or 6 M urea, and sonicated using a MSE sonicator. Sonication was carried out on ice, using six 30 s periods of sonication with 30 s intervals between them. The suspension was then left on ice for between 20 min and 1 h for the urea to dissolve the proteins, then four further 30 s bursts of sonication, with 30 s intervals between them, were applied.

When urea was not being used, the frozen pellets were re-suspended in binding buffer without urea and kept on ice at all times to limit the action of proteases. Sonication was carried out using the same equipment settings but for shorter bursts to prevent heating

the sample: between 6 and 20 x 15 s bursts were used with 30 s intervals in between each burst. Microscopic examination was employed to determine the extent of bacterial disruption.

In both cases, the suspension was then decanted into 1.5 ml microcentrifuge tubes and centrifuged at 14,000 rpm for 30-40 min (in a centrifuge refrigerated to 4°C where urea was not used). The supernatant was removed with a syringe and the end of a pipette tip and filtered through a 5 µm filter ready for protein purification.

### *2.3.2 Protein purification by metal-ion chelation column chromatography*

This procedure was initially carried out using in-house sepharose columns. Later in the course of the project commercially available nickel columns were used (Quick-900 columns, Novagen).

#### *2.3.2.1 Protein purification using in-house sepharose columns*

When urea was used, all the buffers except for the charge buffer and the strip buffer contained 8 M urea. When urea was not used, extraction and purification was performed keeping temperatures at 4°C throughout. The steps up to purification were done on ice and using a refrigerated centrifuge, and the whole purification procedure was carried out in a 4°C cold room.

Columns were prepared by adding 2 ml resin (iminodiacetic acid immobilized on Sephadex 6B Fast Flow, suspended in 20% ethanol- Sigma) to a chromatography column (Sigma). The resin was allowed to settle. The columns were then rinsed twice with 10 ml deionised water, to wash out the alcohol, and then charged with 10 ml charge buffer (containing nickel). When this had eluted, the column was equilibrated

with 10 ml binding buffer. After this the base of the column was blocked and 2 ml binding buffer was added. The surface was gently pipetted up and down to create a flat surface after re-settling (the end of the pipette tip was cut in order not to disturb the beads too much). Then the remaining binding buffer was allowed to drain off.

The filtered supernatant (crude extract) containing the cell protein was then applied to the column, and the unbound fraction collected. About 20  $\mu$ l of the crude extract was also kept aside for later analysis by SDS-PAGE. The column was washed with 20 ml binding buffer. 50 ml wash buffer containing 60 mM imidazole was then applied to the column and a sample of the bulk eluate collected (wash fraction). A few ml of elution buffer were applied and 1 ml fractions were collected in 5 or 6 numbered microcentrifuge tubes (elution fractions). The column was then washed with 8 ml strip buffer and a sample of the bulk eluate collected again (strip fraction). Finally, 10 ml of deionised water was passed through the column, followed by 10 ml of 20% ethanol. The columns were stored at 4°C with 3 ml of 20% ethanol in them.

#### *2.3.2.2 Protein purification using Quick-900 columns (Novagen)*

Later in the course of the project commercially available nickel columns were used (Quick-900 columns, Novagen), instead of in-house sepharose columns. The buffers used contained 6 M rather than 8 M urea and the manufacturer's instructions were followed.

#### *2.3.3 Estimation of purified protein concentration*

The concentration of protein in each of the elution fractions was tested by one of two methods – either by measuring the  $A_{280\text{nm}}$  in a spectrophotometer, against a blank of elution buffer, and using this in conjunction with the molar extinction coefficient

(Appendix *i*), to calculate the concentration; or by using a modified Bradford assay, available as Coomassie Plus Protein Assay kit (Pierce) and following the manufacturer's instructions.

#### 2.3.4 Large scale expression, extraction and purification of *RpfE*

Eight litres of culture of the *E coli* host were set up for induction, in a yeast-peptone induction medium. Ampicillin 50 µg/ml was added, plus 1 ml glucose per 100 ml culture to suppress protein production until IPTG was added.

Sixteen 500 ml flasks were set up. Each was inoculated with a 1:100 dilution of overnight culture of clone AD11 – *rpfE* in vector pET-15b.tev and host Origami B (DE3) pLysS (see Appendix *iii* for a table of clones used). IPTG 0.1 mM was added when the OD<sub>600</sub> was 0.7 – 0.9. The flasks were incubated at 25°C for a further 4 ½ hours. The cultures were spun down and washed with PBS. Extraction was done on the deposit from 2 l culture volumes at once and 2 ml lysis (binding) buffer was used per 1g pellet. 1 tablet of Protease Inhibitor Cocktail (Roche) was added per 50 ml lysis suspension. The suspension was sonicated on ice using 12 bursts of 30 s with 30 s intervals. Sonicated suspension was spun down at 20 000 rpm for 1 h.

The cell extract was loaded onto an AKTA<sup>TM</sup> FPLC<sup>TM</sup> chromatography system (Amersham Biosciences) and run down two 5 ml HiTrap nickel-chelating HP columns<sup>TM</sup> (Amersham Biosciences). Subsequently, the imidazole concentration in the buffers was increased at discrete intervals by mixing binding and elution buffer.

5 ml aliquots which corresponded with protein peaks were saved and kept at 4°C.

Desalting was performed on the same equipment using HiPrep 26/10 50ml desalting columns<sup>TM</sup> (Amersham Biosciences) and binding buffer.

The equipment was supported by the Unicorn V3.00 software operating system.

### 2.3.5 TEV protease cleavage of His-tag from RpfE expressed from pET-15b.tev

Recombinant TEV (rTEV) protease was obtained from Invitrogen in concentration 10 u/μl. The reaction mix consisted of:

1 μl rTEV protease per 20 μg protein.

1% v/v DTT (0.1 M)

0.1% v/v EDTA (0.5 M)

The mixture was incubated on a roller at 16°C. Small aliquots were taken off hourly for the first 3 h and then after overnight incubation. Once removed the reaction was stopped in each aliquot by heating in a hot-block at 80°C for 10 min. Each sample then was then concentrated using Vivaspin<sup>TM</sup> filter-concentrators (Vivascience). rTEV protease was separated from Rpf by passing through a HiTrap<sup>TM</sup> nickel-chelating FPLC column as before.

### 2.3.6 SDS-PAGE (Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis)

The equipment used was the BioRad Mini-PROTEAN 3<sup>TM</sup> system. Acrylamide was obtained from National Diagnostics, TEMED and ammonium persulphate (APS) from Sigma. A 12.5% separating gel was made as follows:

Acrylamide	3.275 ml
Tris-HCl pH 8.8	1.86 ml
H <sub>2</sub> O	2.725 ml
TEMED	15 μl
APS	10 μl



The separating gel was overlaid with H<sub>2</sub>O to obtain a flat surface and allowed to polymerise for at least 20 min. Stacking gel was made as follows (equivalent to a 7.5% gel):

Acrylamide	0.5 ml
Tris-HCl	pH 6.8 0.75 ml
H <sub>2</sub> O	1.7 ml
TEMED	20 µl
APS	15 µl

When the separating gel had polymerised, the water was drained off, the stacking gel was applied, the comb inserted and this was allowed to polymerise for at least another 25 min. When polymerised, the comb was removed from the stacking gel and the gel placed in the gel tank, which was filled with running buffer. The gel was loaded with samples of volume 10-15 µl in sample buffer (25% v/v), which had been boiled for 5 min, and a protein marker (NEB) in one well. The gel was run at 200 V for about 60 min, and then stained by incubating overnight at 48°C with Brilliant Blue G-colloidal stain (Sigma) with 20% methanol.

### 2.3.7 *Immuno-blotting*

SDS-PAGE gels run for the purpose of immuno-blotting were run with a prestained MW marker (NEB). After separating the proteins by SDS-PAGE, the gel was washed with distilled water to remove SDS, and the stacking gel removed and discarded. The BioRad Mini-PROTEAN 3<sup>TM</sup> equipment was set up according to the manufacturer's instructions. From negative to positive, the blotting apparatus consisted of a synthetic porous pad, four pieces of filter paper (Whatman), the SDS-PAGE gel, a cellulose

nitrate transfer membrane (Whatman), four more pieces of filter paper, and another porous pad. All components were first soaked in transfer buffer. Bubbles were removed by rolling a glass pipette over the filter papers. The blotting apparatus was immersed completely in transfer buffer in an electrophoresis tank and run at 20 V in a 4°C cold room overnight.

The membrane was removed carefully, using gloves and tweezers, and placed in blocking buffer at 37°C for 30 min. This buffer was tipped off, and fresh blocking buffer containing the primary antibody was added. The membrane was incubated for 2 h at 37°C, then washed with TBS containing 0.2% Tween 20 (3 washes of 5-10 min each, on a 'belly dancer' shaker). Blocking buffer containing 1/20,000 dilution of the secondary antibody was added and the membrane incubated for 30 min at 37°C, before washing again in TBS three times. A SIGMA FAST 5-Bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazodium (BCIP/NBT) tablet (Sigma) was dissolved in the dark in 10 ml distilled water and added to the membrane. This was incubated in the dark until the bands became visible (about 10-20 min) and then the reaction was stopped by discarding the solution and adding water.

## **2.4 Establishing dormant cultures of *Mycobacterium bovis* BCG**

The strain of *M. bovis* BCG used was the Glaxo strain (a derivative of the Pasteur strain). Two types of culture were set up: glass test tubes with air-tight PTFE seals, and 500 ml glass Erlenmeyer flasks with airtight rubber bungs as seals.

*M. bovis* BCG was cultured in Sauton's medium. Differences in viability between cells are more readily apparent in this minimal medium. Use of a rich medium may mask the

effect of addition of Rpf to cultures. In both the tube and the flask models, the vessels contained one fifth of their volume of Sauton's medium, supplemented with ADC.

The same culture of BCG was used to inoculate both the tubes and the flasks. This was a six-day old turbid culture, grown from a frozen stock in Sauton's medium at 37°C.

#### *2.4.1 Tubes*

##### *2.4.1.1 Tubes— culture and sampling*

A final dilution of  $10^{-6}$  of the BCG culture in Sauton's medium was prepared, calculated to contain approximately 2 cfu/ml of organisms. 3 ml aliquots were distributed into tubes. The tubes were incubated at 37°C in a rack and were not shaken. The pellet of BCG was visible at the bottom of the tube from a few days onwards. Each time a sample was taken, a tube was shaken to suspend the cells, opened, the sample taken and the rest of the culture discarded.

##### *2.4.1.2 Tubes – estimation of viable count*

At various intervals, samples were taken and serially diluted in Sauton's medium. Samples (100 µl) of each dilution were spread on agar plates containing either Sauton's medium or Middlebrook medium. Colony counts were performed in order to obtain an estimate of the viable cfu/ml at different times.

#### *2.4.2 Flasks*

##### *2.4.2.1 Flasks – culture and sampling*

A  $10^{-3}$  dilution of the 6 day-old BCG culture in Sauton's medium was inoculated into 100 ml fresh Sauton's medium in a 500 ml Erlenmeyer flask, to give a final dilution of  $10^{-5}$ . The resulting culture was calculated to contain approximately 20 cfu/ml. The flask was incubated on a shaking platform, at 70 rpm and 37°C. Sampling was done after

wiping the top of the rubber septum with alcohol, using a sterile needle inserted through the septum.

#### 2.4.2.2 *Flasks – estimation of viable count*

The sample was serially diluted in Sauton's medium and 100 µl of each dilution was spread on agar plates containing either Sauton's medium or Middlebrook medium. Colony counts were performed in order to obtain an estimate of the viable cfu/ml at different times. Estimation of cfu/ml was carried out in the same way on day 0 to determine the starting inoculum.

#### 2.4.2.3 *Flasks – estimation of Most Probable Number*

On some occasions MPN tests were also performed. This test gives an indication of the number of cells present which can grow in liquid medium (not necessarily the same as the number that can grow when plated directly on agar). MPN is recognised to give a higher viable count than plating on agar. Additionally, Rpf could be added to the liquid medium to give counts both in the presence and in the absence of Rpf, whereas it has no activity when incorporated into agar. Where there was a higher count using Rpf than without, dormant but rescusitatable cells were considered to be present. MPN determinations were performed in Sauton's medium as described in 2.5.2.2

## 2.5 Protein Activity Testing

### 2.5.1 *Protein activity testing measuring optical density of M. luteus cultures*

As *M. luteus* grows far more rapidly than either *M. bovis* BCG or *M. tuberculosis*, attempts were made to use the effect of Rpf on the optical density (OD) of *M. luteus* cultures to assess its activity. *M. luteus* strain NCIMB 13267 ('Fleming strain' 2665) was used. *M. luteus* was grown overnight in 10 ml Nutrient Broth E at 30°C with continuous shaking. This provided a culture containing approximately  $10^9$  cfu/ml. Cells

were collected from 1 ml by centrifugation and resuspended in 1 ml Lactate Minimal Medium (LMM).

The rest of the procedure was carried out in a laminar flow cabinet to avoid contamination. The suspension was passed through a needle 10-15 times to break up clumps of cells and then serially diluted in microcentrifuge tubes to obtain a suspension of about  $10^5$  cfu/ml.

Freshly prepared Rpf was diluted 100-fold in LMM. It was then passed through a 0.2  $\mu$ m filter that had been washed first with 20 ml LMM. Serial dilutions of the filter-sterilised Rpf were then made, to obtain  $\mu$ M- fM concentrations.

In a microtitre plate, wells were inoculated with 145  $\mu$ l of each dilution of Rpf from  $\mu$ M to fM. Six replicate rows were inoculated. The wells around the edge of the plate were filled with LMM alone, to combat the 'edge effect' of evaporation. To each of the other wells were added 5  $\mu$ l of the *M. luteus* cell suspension. The final concentration of *M. luteus* in the wells was  $\sim 3 \times 10^3$  cfu/ml. A control microtitre plate was inoculated in parallel, using elution buffer instead of Rpf. The plates were incubated at 30°C on a shaker and read daily, up to 5 days, looking initially for a deposit of cell aggregates in wells, which is usually followed by development of a turbid cell suspension, indicating growth, in the wells.

#### 2.5.2 Protein activity testing using *M.bovis* BCG cultures

Evidence of activity of fresh Rpf was sought either by performing an activity test, analogous to the one used for *M. luteus*, using different concentrations of Rpf on dormant cultures of BCG, or by carrying out Most Probable Number (MPN)

experiments on dormant cultures both with and without Rpf. The former method has the advantage that the most active concentration of Rpf can be determined. The latter method is easier to interpret as it has a better-defined end-point, and data for plotting growth curves can be obtained, but owing to the numbers of tubes required and space limitations in the incubator, only a few different protein concentrations/samples can be tested at any one time. All testing of activity on BCG was carried out in a Class 2 safety cabinet in a Containment Level 2 laboratory. Details of how dormant BCG cultures were obtained are described in 2.4.

#### 2.5.2.1 'Concentration-activity test' method

Dormant *M. bovis* BCG cells were used. 1 ml of a  $10^{-2}$  dilution of the cells was inoculated into each of 6 tubes containing 9 ml Sauton's medium, thus producing a dilution of  $10^{-3}$  in each tube.

Freshly prepared Rpf was sterilised by passing a 100-fold dilution in Sauton's medium through a 0.2  $\mu\text{m}$  filter, which had been flushed with 20 ml Sauton's medium. It was then serially diluted, and different dilutions were added to 5 of the 6 tubes of BCG. No Rpf was added to the 6<sup>th</sup> tube, which acted as a control. The final dilutions of Rpf ranged from nM to fM. The contents of each tube were then mixed and pipetted into a rack of tubes in 2 ml aliquots, to provide 5 replicates of each concentration of Rpf, and the controls. The tubes were incubated at 37°C. They were read daily and tubes in which visible growth was seen were recorded as positive. Growth appeared as a white pellet at the bottom of the transparent tube. Results were recorded daily until all the tubes contained visible growth.

### 2.5.2.2 *Most Probable Number method*

For this test, different dilutions of dormant BCG cultures were used with the same dilution of Rpf in each tube (as opposed to different concentrations of Rpf in the same dilution of BCG, as described in 2.5.2.1). The MPN tests were performed in Sauton's medium as follows:

‘Master-mix’:

$10^{-3}$  dilution: 22.5 ml Sauton's medium + 2.5 ml dormant BCG culture at  $10^{-2}$  dilution

$10^{-4}$     “       “       “       “       “       “       at  $10^{-3}$  dilution

$10^{-5}$     “       “       “       “       “       “       at  $10^{-4}$  dilution

$10^{-6}$     “       “       “       “       “       “       at  $10^{-5}$  dilution

$10^{-7}$     “       “       “       “       “       “       at  $10^{-6}$  dilution

$10^{-8}$     “       “       “       “       “       “       at  $10^{-7}$  dilution

Samples (2 ml) of each ‘Master-mix’ were incubated in 5 replicate clear plastic tubes at 37°C for up to 6 weeks. These contained no Rpf. To the remaining 12.5 ml of each ‘Master-mix’ were added 0.125 ml of the appropriate dilution of filter-sterilised Rpf. The required dilution of Rpf depended on the calculated concentration of the stock. Normally a final concentration of ~30 pM was aimed for in the MPN experiments. Samples (2 ml) of each of the ‘Master-mixes’ + Rpf were incubated in 5 replicate clear plastic tubes along with the controls. The tubes were all incubated at 37°C and read daily until no additional tubes showed new growth over a period of 10 days. This was taken as the end point of the experiment. The results were interpreted using published MPN tables (Appendix v).

### 2.5.3 Protein activity testing using ATP-bioluminescence of *M. luteus* cultures

#### 2.5.3.1 Overnight *M. luteus* cultures ('overnight test')

This technique was applied to assay changes in bacterial counts in the presence and absence of recombinant Rpf. As a positive control, Rpf in the supernatant of stationary-phase *M. luteus* cultures was used.

LMM was inoculated with an overnight culture of *M. luteus* ( $OD_{600} = 2.0$ ) in nutrient broth to give a thousand-fold dilution (equivalent to  $\sim 10^6$  cfu/ml in a 50 ml total volume). This was divided into 4 ml aliquots. To each aliquot was added 40  $\mu$ l freshly-produced Rpf to give dilutions of Rpf between  $\mu$ M and pM. For the negative control, an equal volume of a  $10^{-3}$  dilution of elution buffer was used. For the positive control, 2 ml sterile filtered supernatant of a stationary phase culture of *M. luteus* in LMM was added to 2 ml of the inoculated LMM. Each tube was set up in duplicate. The tubes were incubated overnight at 37°C on a shaker-incubator. After 16-18 h the tubes were removed and the cultures spun down. The pellet was resuspended in 50  $\mu$ l of 0.1 M Tris/0.004M EDTA buffer, pH 7.75. This suspension was added to a boiling 450  $\mu$ l volume of the same buffer. The sample was then heated to 100°C again for 2 min, then spun down at 1000g for 1 min. The supernatant was removed and kept on ice prior to measurement of ATP-bioluminescence in a luminometer, using the ATP Bioluminescence Assay Kit CLS II (Roche) as per the manufacturer's instructions.

ATP concentration was calculated as follows: for each experiment, using ATP standards a standard curve was drawn, on a log-log graph, to correlate bioluminescence with ATP concentration. The ATP concentrations in the cultures could then be read off the graph.



#### 2.5.3.2 *Four hour M. luteus cultures ('rapid test')*

A modification of the test in 2.5.3.1 was developed in order to get an activity test result on the day of production of the protein, rather than the day after. In this more rapid test, the initial overnight culture of *M. luteus* ( $OD_{600} \sim 2.0$ ) was the same, but instead of diluting it 1000-fold it was only diluted down to an  $OD_{600}$  of 0.1. The diluted cells were then aliquoted as described in 2.5.3.1 but only cultured for 4 h at 37°C before ATP was extracted.

## 2.6 Phage display

### 2.6.1 *Construction of phage display library*

Since a relatively large quantity of *M. tuberculosis* DNA would be required to construct a phage library, broth cultures were set up. 200 ml of 7H9 broth was inoculated with a 2 ml culture of *M. tuberculosis* H37Rv, and incubated at 37°C for 6 weeks. The culture was spun down and DNA extraction was carried out as described in 2.1.2.

Aliquots of the DNA were cut with the blunt-ended restriction enzymes *Hae*III, *Rsa*I, *A*luI, and *H*pyCH4U (NEB). To obtain DNA fragments in the size range required, digests were done using various concentrations of DNA and enzyme, and digestion was continued for various time periods. The aim was to find a concentration of DNA, restriction enzymes and lengths of digestion times which would provide a library of fragments between ~1.5 kb and ~200 bp in size. Digests were stopped after various lengths of time by heating to 80°C for 20 min. 2-3 µl of each partial digest were then run on a 1% TAE agarose gel to assess the size of the fragments in each. The protocol found to be successful was as follows:

Extracted DNA was diluted in dH<sub>2</sub>O 1:4 v/v. 1 µl of each of the four enzymes was added to 26 µl dH<sub>2</sub>O. The digest reaction was then performed as follows:

2 µl enzyme mixture (diluted 1 µl of each enzyme in 26 µl dH<sub>2</sub>O)

15 µl DNA solution

2 µl restriction buffer

1 µl dH<sub>2</sub>O

(20 µl total volume)

The reaction was incubated at 37°C. 3 µl aliquots were removed at intervals of 2, 5, 15 and 35 min, heated to 80°C for 20 min to stop the reaction, and then run on a 1% TAE to check that the size range of the DNA fragments was as required. To make the library, a larger number and volume of DNA fragments was necessary, so the reaction was scaled up tenfold. From this reaction, the 25 min, 35 min and 60 min fractions were pooled and then cleaned using the Qiagen gel extraction kit.

Plasmid pG8H6 was supplied by Brian Henderson, Eastman Dental Institute. The plasmid was cut using *Sma*I (NEB) to give blunt ends, and dephosphorylated with calf intestinal alkaline phosphatase (NEB) to prevent self-ligation of the vector in subsequent ligations. The *M. tuberculosis* DNA fragments were then ligated into the cut plasmid pG8H6 using the Ready-To-Go<sup>TM</sup> T4 DNA ligase system (Amersham Biosciences). The reaction used was:

13 µl digested DNA

3 µl cut plasmid DNA

4 µl dH<sub>2</sub>O

The ligated DNA was cleaned with the Qiagen Minelute kit and resuspended in 10 µl dH<sub>2</sub>O.

Electrocompetent *E coli* TG1 cells were obtained from Stratagene. 2 µl of the ligation product was added to 100 µl of these cells and transformed by electroporation using a Bio-Rad electroporator. A 0.1 cm gap cuvette was used and a voltage of 1700 V.

900 µl of nutrient broth warmed to 37°C was added immediately and the cells incubated on a shaker-incubator at 37°C for 1 h. 10 µl and 20 µl aliquots were then plated onto nutrient agar containing 100 µg/ml ampicillin and incubated overnight at 37°C. Resultant clones were sub-cultured in nutrient broth containing 100 µg/ml ampicillin overnight. Plasmid DNA was extracted from each using the Qiagen Miniprep kit. The DNA was then digested with *HindIII* and *PstI* to check for the presence and size of inserts in the plasmid. Three electrotransformation reactions were performed and the transformants pooled.

#### *2.6.2 Transfection and PEG precipitation of phage*

The cells in the library were transfected with helper phage R408 (Promega). The cells were warmed to 37°C for a few minutes, then 200 µl of phage was added to 1 ml of library (a 20-fold excess of phage over cells). The mixture was left to stand at 37°C for 15 min and then added to 100 ml nutrient broth containing 100 µg/ml of ampicillin. The broth was cultured overnight at 37°C on a shaker-incubator. The following morning the culture was spun down for 10 min at 5000. 14.3 ml of 25% PEG 6000 in 2.5 M NaCl was added to the 100 ml of supernatant, split between 4 Falcon tubes. The mixture was mixed gently and left at 4°C overnight. The precipitated phage was then spun down at 5000 rpm for 15 min. A translucent gel-like precipitate was visible coating the sides of the Falcon tubes. The supernatant was discarded carefully, and the precipitate resuspended in 5 ml TE buffer and stored at 4°C.

### 2.6.3 *Titering of phage stocks and fractions from panning*

10 ml of nutrient broth was inoculated with 50 µl of an overnight culture of TG1 cells and grown until the OD<sub>600</sub> was 0.5 (3 h). To titre phage stocks, 10 µl of 10<sup>-5</sup> – 10<sup>-8</sup> dilutions of the phage stock was added to 200 µl of bacteria from this culture.

To titre fractions collected from panning, 100 µl of appropriately diluted phage was added to 400 µl of bacteria:

Unbound: 10<sup>-5</sup> – 10<sup>-8</sup> dilutions

Last wash: neat – 10<sup>-2</sup> dilutions

Bound: neat – 10<sup>-4</sup> dilutions

The mixture was incubated for 30 min at 37°C. 100 µl volumes were then plated out on nutrient agar plates containing 100 µl/ml ampicillin. Colony counts were performed the following day after overnight incubation at 37°C.

### 2.6.4 *Panning for RpfB receptor protein*

Fresh recombinant RpfB was used. The protein was first dialysed in PBS using a 0.5-3.0 ml Slide-A-Lyzer<sup>TM</sup> dialysis cassette (Pierce) to remove urea and imidazole. A

5 ml polystyrene tube was then coated with the protein overnight at 4°C. The following day RpfB was removed from the tube which was then blocked for 2 h with PBS 2% BSA. BSA was also added to the library prior to the panning to avoid selection for any phage with affinity for it. The polystyrene tube was washed with PBS and 1ml of the library was added to the tube. The tube was left rotating for 2 h at room temperature (or 4 h in the first round of panning only). The library was then removed from the tube and kept (unbound fraction) for colony counts. The tube was then washed with 10 x 4 ml of PBS 0.05% Tween, then 10 x 4 ml of PBS. The final wash was kept. The bound phage was removed from the tube by adding 1 ml 0.1 M glycine pH 2.1 for 10 min at room

temperature. The glycine was then removed and added to 0.5 ml 1 M Tris pH8.0. The phage fractions were then titred according to the protocol in 2.6.3.

#### 2.6.5 Sequencing and analysis of inserts from clones selected by panning

To investigate the clones selected after three rounds of panning, the inserts were sequenced. The primers used to amplify the inserts from pG8H6 had the following sequences: 5'-TTG CCT ACG GCA GCC GCT GAA-3' and 5'-TGC GGC CCC ATT CAG ATC CTC-3' (Jacobsson 1995, Heilmann 2002). Sequencing was performed on a thermal cycle sequencer.

#### 2.6.6 Cloning and expression of Rv0328

Rv0328 was amplified by PCR from *M. tuberculosis* genomic DNA using the primers and protocol described in 2.1.4; cloning into the TOPO<sup>TM</sup> vector and DH5 $\alpha$  host, and sub-cloning into pET-21a and Origami DE3 were also done as described previously in 2.1.

The protein was expressed as described in 2.2, with a 3 h induction with IPTG. Purification was with Novagen Quick 900 columns.

### 2.7 Immunocytochemistry of *M. tuberculosis*-infected tissues with anti-Rpf antibodies

Purified polyclonal sheep anti-Rpf antibodies were provided by Mike Young (Mukamolova *et al* 2002b). Secondary rabbit anti-sheep antibodies, and tertiary antibodies (peroxidase-conjugated streptavidin) were obtained from Jackson ImmunoResearch. Tissue infected with *M. tuberculosis*, and all control tissue, was provided by the histopathology department at the Royal Free Hospital.

### *2.7.1 Immunocytochemistry protocol*

Sections were deparaffinised in xylene for 10 min, then rinsed twice in clean xylene for a few seconds each, followed by 3 rinses in industrial methylated spirits to remove the xylene. Endogenous peroxidase activity was blocked by immersing the sections in blocking agent (590 ml methanol: 10 ml water) for 10 min. The slides were then washed under running water for 5 min. It was found that pre-treatment of samples by enzyme digestion or heat was not required. The slides were transferred to a humidity chamber and the sections ringed with a hydrophobic pen. Non-specific protein binding was blocked by covering the sections with normal rabbit serum (NRS) (Stratech) at a 1:10 dilution (equivalent to 6 mg/ml) for 20 min at room temperature. The NRS was shaken off and the sections covered with primary antibody at an appropriate dilution, and left to stand in the humidity chamber for 1 h. A negative control was always used in which the NRS was left on the slide and the addition of primary antibody was omitted.

The sections were then washed with TBS which was left on for 6 min, before shaking off and applying secondary antibody for 30 min. Then the sections were washed again with TBS as previously (left on for 6 min) and peroxidase-conjugated streptavidin (tertiary antibody) at a concentration of 1 µg/ml was added for 30 min. After a further wash with TBS the sections were developed with freshly prepared diaminobenzidine (DAB) solution (Vector) for 8 min. They were then washed again with TBS and then running water for 5 min. Counter-staining was performed with Mayer's haematoxylin for 5 min, then sections were rinsed in water again before being dehydrated by dipping in alcohol and then cleared by dipping in xylene. Slides were mounted in a synthetic mountant (DPX).

The correct working dilution of rabbit anti-sheep secondary antibodies was determined by binding various dilutions of the antibodies to sheep anti-factor 8 (Serotec) bound to sections of human placenta. The secondary antibody was tested at dilutions of 1:1000 and 1:5000, against dilutions of the sheep anti-factor 8 antibody of 1:100, 1:500, 1:1000 and negative control (no primary antibody).

Next, to determine the optimum dilution of anti-Rpf antibodies, a dilution of 1:1000 rabbit anti-sheep (secondary) antibody was used against the following dilutions of anti-Rpf (primary) antibodies: 1:100, 1:500, 1:1000, 1:2000. Experiments using the anti-Rpf antibody were performed on human tissue samples known to contain *M. tuberculosis* (from gastro-intestinal lymph nodes).

#### *2.7.2 Absorption of anti-Rpf antibodies using RpfB*

These experiments were carried out in order to demonstrate that positive immunohistochemistry results were due to binding of the anti-Rpf antibodies to Rpf. The addition of recombinant RpfB to the anti-Rpf antibodies at the start of the experiment should neutralise them and lead to a negative immunohistochemistry result. Fresh recombinant RpfB was added to the anti-Rpf antibodies and left at room temperature for 1 hour. Then the staining was carried out as described in 2.7.1.

#### *2.7.3 Ziehl-Neelson staining of sections*

In some immunohistochemistry experiments, sections were stained by the cold Ziehl-Neelson technique after to immuno-staining, to demonstrate the location of acid-fast bacilli in the sections. Slides were left to stand in filtered carbolfuchsin (BDH fuchsin RAL 22 Cold) in a Coplin jar for 10 min. After washing in tap water they were differentiated by dipping several times in 1% acid alcohol. This was repeated until the

sections were a very pale pink colour, then they were washed in tap water for 2 min. Counterstaining was performed with Carazzi's haematoxylin for 3 min, then the slides were washed again in running tap water for 5 min. Slides were then dehydrated, cleared and mounted.

#### *2.7.4 Staining of sections with anti-lipoarabinomannan antibodies*

Anti-lipoarabinomannan (LAM) antibodies were obtained from Vladimir Koulchin's laboratory at Chemogen, Portland USA. The antibodies were raised in rabbits and conjugated to horseradish peroxidase. They were used at a 1:50 dilution, with peroxidase-conjugated streptavidin as used in the anti-Rpf experiments. The method used was as described in 2.7.1.



## **CHAPTER III**

### **RESULTS: CLONING AND EXPRESSION OF *M.TUBERCULOSIS***

**Rpfs**

### 3.1 Introduction

#### 3.1.1 Aims

The aim of the work described in this chapter was to express the cloned genes for the five *M. tuberculosis* RpfS and to purify the proteins for use in subsequent experiments. RpfD was cloned *de novo* from *M. tuberculosis* DNA into the pET-21a (Novagen) vector and expressed in the *E. coli* host Origami B (DE3) pLysS (Novagen). This host and vector had not previously been used for the expression of RpfS. The other RpfS were initially expressed in vector pET-19b (Novagen) and *E. coli* host HMS174(DE3) (as provided by Mike Young's group, University of Wales, Aberystwyth) and then subcloned into expression host Origami B (DE3) pLysS, with the aim of achieving better solubility of the expressed protein.

#### 3.1.2 Host strain

Origami B (DE3) pLysS is a strain with mutations in both thioredoxin reductase (*trxB*) and glutathione reductase (*gor*), which permits the formation of disulphide bonds in the cytoplasm and thus may enable optimal folding of a protein that is normally secreted. Since the RpfS contain cysteine residues it was thought that disulphide bond formation between these residues might play an important part in the activity of the proteins. Studies have shown that expression in Origami (DE3) yielded 10-fold more active protein than in another host even though overall expression levels were similar (Prinz 1997). In addition the pLysS plasmid provides T7 lysozyme to reduce basal expression of target genes and therefore stabilise plasmids that express proteins toxic to *E. coli*. All Origami B strains contain a *lacY* deletion mutation which allows control over the amount of protein expression by controlling the concentration of the IPTG inducer.

### 3.1.3 Vectors

The pET system for the cloning and expression of recombinant proteins in *E. coli* has been developed by Novagen. Target genes are cloned in pET plasmids under control of strong bacteriophage T7 transcription and translation signals: expression is induced by providing a source of T7 RNA polymerase in the host cell. T7 RNA polymerase is so selective and active that almost all the cell's resources are converted to target gene expression; the desired product can comprise more than 50% of the total cell protein after a few hours of induction. The system maintains target genes transcriptionally silent in the uninduced state. Target genes are initially cloned using hosts that do not contain the T7 RNA polymerase gene, so that they are switched 'off' and cannot cause plasmid instability due to the production of proteins potentially toxic to the host cell. Once established, plasmids are transferred into expression hosts containing a chromosomal copy of T7 polymerase. This is done by using a lysogen of bacteriophage DE3, which carries a DNA fragment containing the *lacI* gene, the *lacUV5* promoter and the gene for T7 RNA polymerase. The *lacUV5* promoter is inducible by the addition of isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Different pET vectors have different features which may be useful in different settings. For example, pET-19b has an N-terminal histidine tag and an enterokinase protease site, whilst pET-21a has a C-terminal histidine tag and no protease cleavage site.

pLysS is not a cloning vector, but is used to suppress basal expression from the T7 promoter by producing T7 lysozyme, a natural inhibitor of T7 RNA polymerase. The presence of the pLysS plasmid increases the tolerance of the host cell for plasmids with toxic inserts: unstable plasmids become more stable.

Maps of plasmids pET-19b, pET-21a and pLysS, and the cloning and expression regions of pET-19b and pET-21a are shown in Appendix iv.

### 3.2 Materials and Methods

Detailed technical descriptions of the materials and methods are to be found in Chapter 2 sections 2.1 (cloning methods) and 2.2 (protein expression). Here the procedures used are summarised.

#### 3.2.1 Cloning of *rpfD* into pET-21a vector and Origami B (DE3) pLysS host

RpfD is one of the smaller *M. tuberculosis* Rpf (154 aa, predicted molecular weight 15,664 Da) and was chosen for cloning as its size made it more suitable for structural analysis by NMR, at a later date.

DNA was extracted from *M. tuberculosis* H37Rv. PicoGreen<sup>TM</sup> analysis showed that the concentration of the extracted DNA was 50 µg/ml. Primers were designed to amplify the *rpfD* gene without the N-terminal signal sequence and their sequences are shown in 2.1.4. The primers had an *Nde*I site at the N-terminal end and an *Xho*I site at the C-terminal end. The sequence for amplification had a length of 315 bp.

The PCR reaction was set up as described in 2.1.4. The amplified DNA, which was of the required size of 300 bp on agarose gel electrophoresis, was excised from the gel and extracted using the QIAquick gel extraction kit. The DNA was ligated into TOPO<sup>TM</sup>, transformed into DH5α cells, and plated out for blue-white selection.

Approximately 50% of the resulting colonies were white and ten of these were picked off and inoculated into nutrient broth containing ampicillin 50 µg/ml for overnight

culture. Plasmid DNA was extracted from the broth cultures and restriction digests, using *Xho*I and *Nde*I, were done to confirm the presence of a ~315 bp fragment. All 10 clones contained the required fragment.

Two clones were selected, designated AD1 and AD2. The presumed *rpfD* gene inserts from AD1 and AD2 were excised from an agarose gel and purified. To check whether they contained any mutations they were sequenced using laser fluorescence detection (see 2.1.14) and the data was analysed using Chromas software. This revealed that AD1 contained a point mutation near the 3' end but that the sequence of AD2 was correct. Clone AD1 was therefore discarded.

Clone AD2 consisted of the TOPO<sup>TM</sup> vector containing the *rpfD* gene insert in DH5 $\alpha$ . pET-21a (Novagen) was used as the expression vector. pET-21a was digested with *Xho*I and *Nde*I, and the linear plasmid was purified by agarose gel electrophoresis and gel isolation. The *rpfD* gene insert from AD2 was ligated into the digested pET-21a vector.

The product of the ligation reaction was transformed into DH5 $\alpha$  and plated onto ampicillin selective plates. Initial attempts at transformation were unsuccessful, with all resulting clones lacking the insert. Subsequently the ligation was repeated after first treating the digested pET-21a vector with calf-intestinal alkaline phosphatase. One colony was obtained. Digestion of the plasmid present in this colony revealed the presence of a ~ 300 bp insert. This clone therefore consisted of pET-21a vector with *rpfD* insert in the DH5 $\alpha$  host (AD3). The plasmid with the *RpfD* insert was purified and transformed into the commercial expression host, *E. coli* Origami B (DE3) pLysS (Novagen).

The success of the transformation was checked by extracting the plasmid DNA, digesting with *Xho*I and *Nde*I, and running on an agarose gel to visualise the insert. This clone was designated AD4.

### 3.2.2 Transformation of *Origami B (DE3) pLysS* host cells with *rpfA*, *rpfB*, *rpfC* and *rpfE*.

Mike Young's group at University of Wales, Aberystwyth, provided cloned sequences for the other four *M. tuberculosis rpf* genes. The DNA sequences were located in pET-19b vectors in host *E. coli* strain HMS174. However, the proteins were very poorly soluble, requiring a denaturant (urea) to solubilise them, a problem discussed further in Chapter 4. To try and improve the folding, and hence solubility and activity of the expressed Rpf, plasmid DNA was extracted from each of these strains, and transformed into competent *E. coli* *Origami (DE3) pLysS*, as has already been described for the *rpfD* construct (section 3.2.1). These clones in *Origami (DE3) pLysS* were designated AD7 (containing *rpfA*), AD5 (containing *rpfB*), AD6 (containing *rpfC*), and AD8 (containing *rpfE*). They were used later in the course of the project when trying to improve protein solubility and the level of expression (Chapter 4). A table showing the clones constructed during the course of this project is shown in Appendix *iii*.

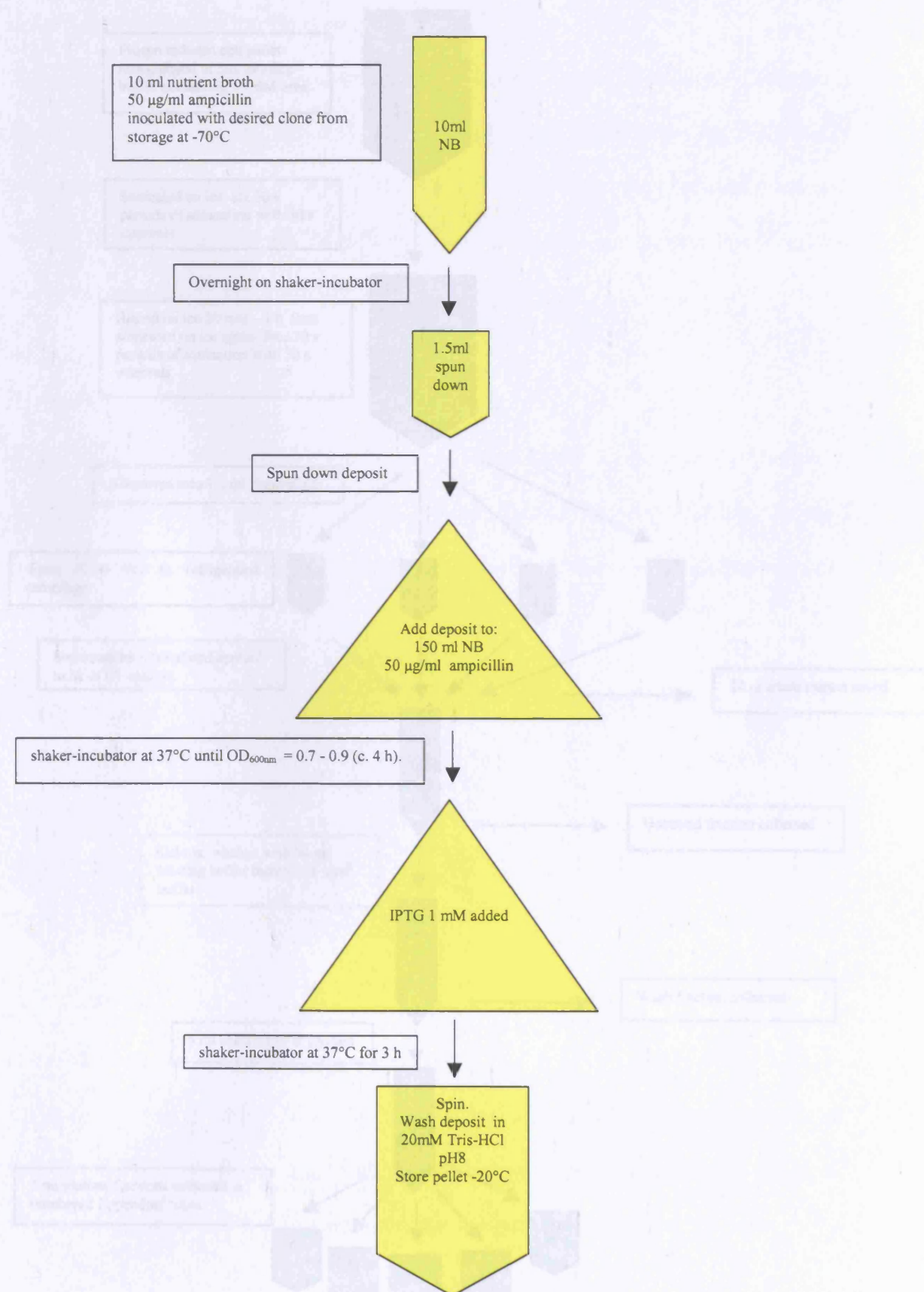
### 3.2.3 Expression and purification of the five *M. tuberculosis Rpf*s

Initially expression was performed from the clones supplied by Mike Young's group at the University of Wales, Aberystwyth. As described in the introduction to this chapter, these consisted of the *rpf* genes (without signal sequence) in pET-19b vectors and HMS174 *E. coli* hosts. The proteins were all extracted, expressed and purified as described in 2.2 and 2.3, using in-house Sepharose columns and buffers containing

8 M urea. Later in the course of the project, for greater convenience, Quick-900 columns (Novagen) were used with buffers containing 6 M urea. The methods used are summarised in the flow charts (Figures 3.1 and 3.2).

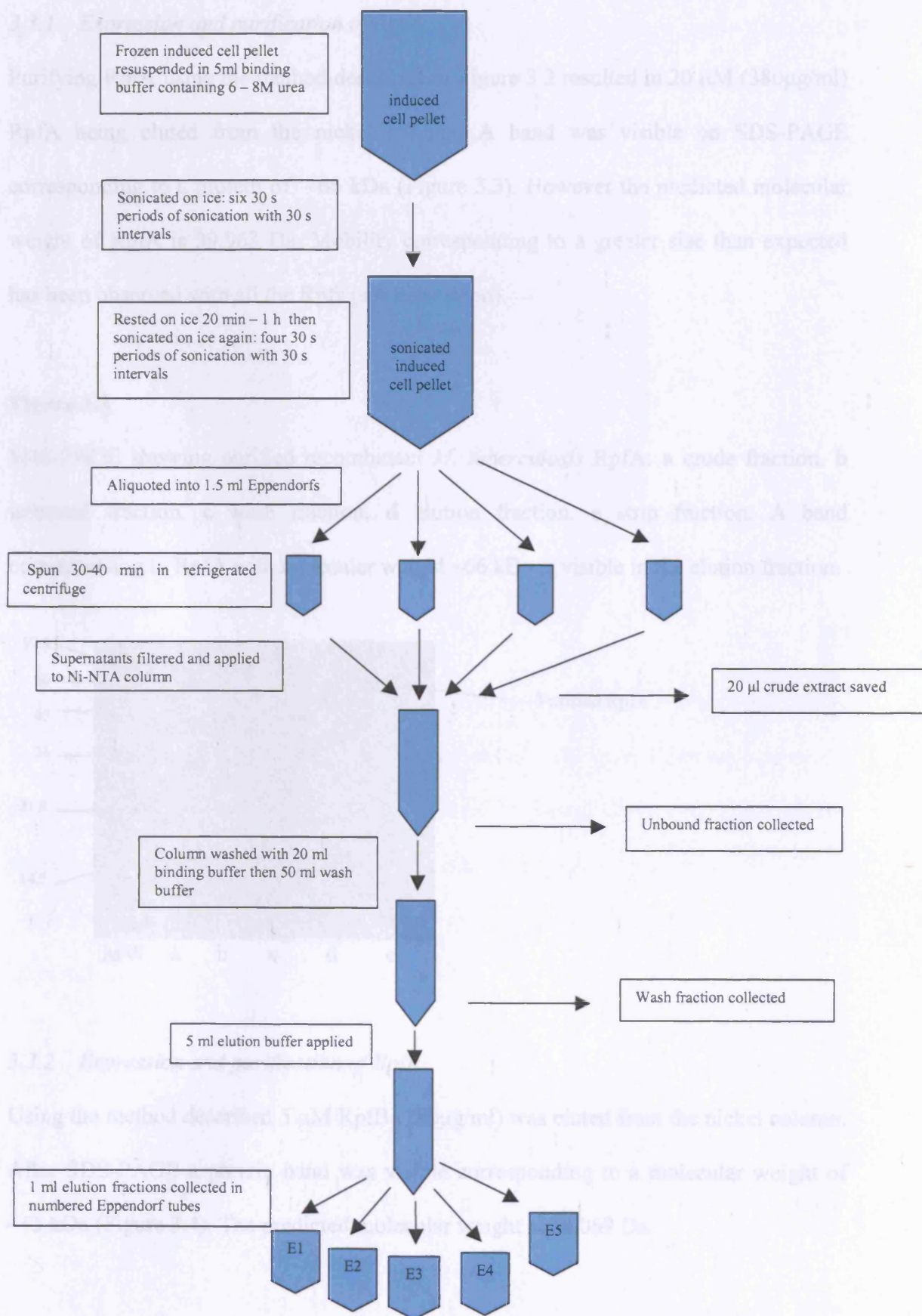
The concentration of Rpf in the purified eluted fractions was calculated using the  $A_{280\text{nm}}$  of the fractions and the molar extinction co-efficient of the Rpf in question (Appendix *i*).

**Figure 3.1** Flowchart showing procedure for induction and expression of Rpfs





**Figure 3.2** Flowchart showing procedure for extraction and purification of RpfS



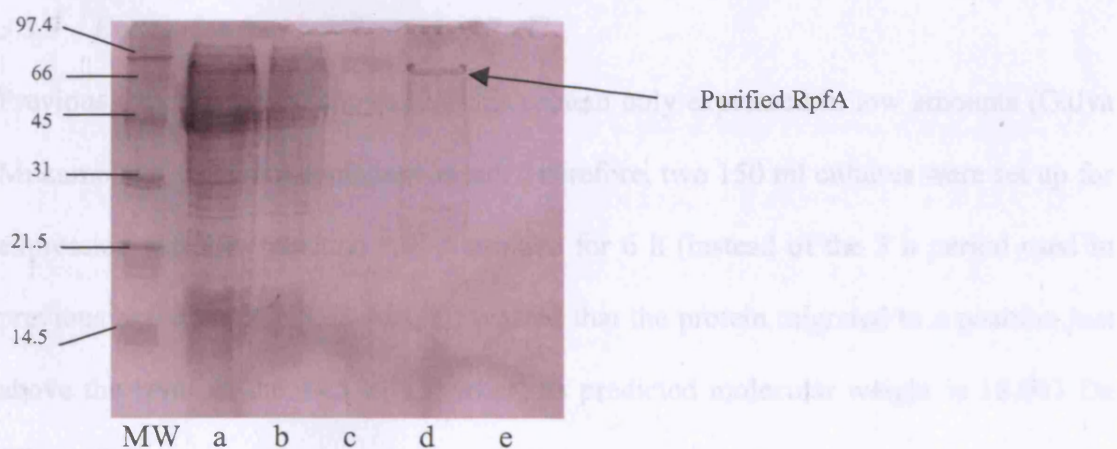
### 3.3 Results

#### 3.3.1 Expression and purification of RpfA

Purifying RpfA using the method described in Figure 3.2 resulted in 20  $\mu$ M (380 $\mu$ g/ml) RpfA being eluted from the nickel column. A band was visible on SDS-PAGE corresponding to a protein of  $\sim$ 66 kDa (Figure 3.3). However the predicted molecular weight of RpfA is 39,963 Da. Mobility corresponding to a greater size than expected has been observed with all the RpfA (see discussion).

**Figure 3.3**

SDS-PAGE showing purified recombinant *M. tuberculosis* RpfA: **a** crude fraction. **b** unbound fraction. **c** wash fraction. **d** elution fraction. **e** strip fraction. A band corresponding to RpfA with molecular weight  $\sim$ 66 kDa is visible in the elution fraction.



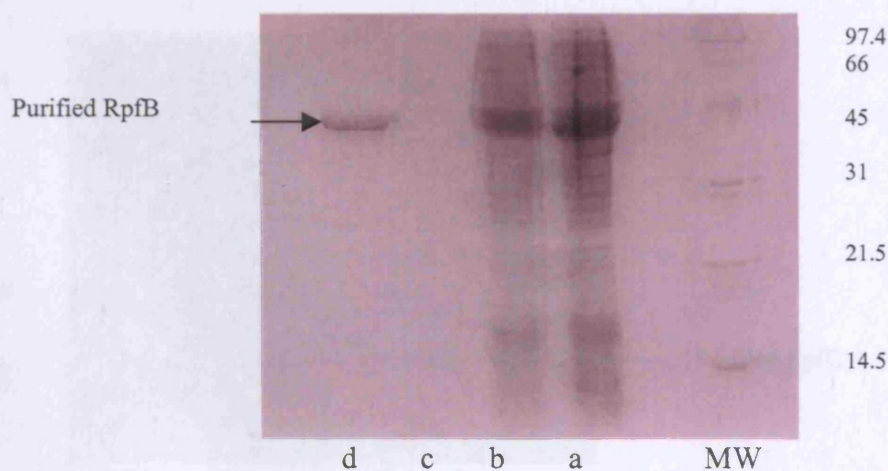
#### 3.3.2 Expression and purification of RpfB

Using the method described 5  $\mu$ M RpfB (180 $\mu$ g/ml) was eluted from the nickel column. After SDS-PAGE a protein band was visible corresponding to a molecular weight of  $\sim$ 45 kDa (Figure 3.4). The predicted molecular weight is 38,069 Da.



**Figure 3.4**

SDS-PAGE showing purified recombinant *M. tuberculosis* RpfB: **a** crude fraction. **b** unbound fraction. **c** wash fraction. **d** elution fraction. A band corresponding to RpfB with molecular weight ~45 kDa is visible in the elution fraction.

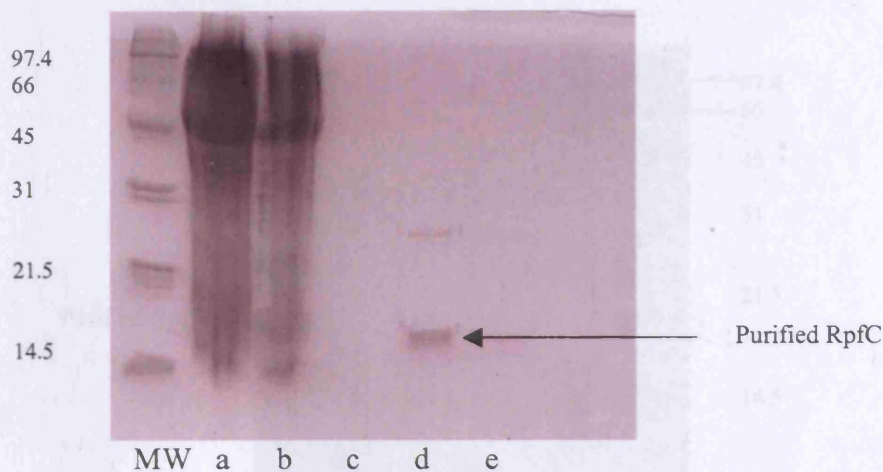


### 3.3.3 Expression and purification of RpfC

Previous experience had shown that this protein only expressed in low amounts (Galya Mukamolova, personal communication). Therefore, two 150 ml cultures were set up for expression and the induction was continued for 6 h (instead of the 3 h period used in previous experiments). SDS-PAGE revealed that the protein migrated to a position just above the level of the 14.5 kDa marker. Its predicted molecular weight is 18,003 Da (Figure 3.5).

**Figure 3.5**

SDS-PAGE showing purified recombinant *M. tuberculosis* RpfC: **a** crude fraction. **b** unbound fraction. **c** wash fraction. **d** elution fraction. **e** strip fraction. A band corresponding to RpfC is visible in the elution fraction.

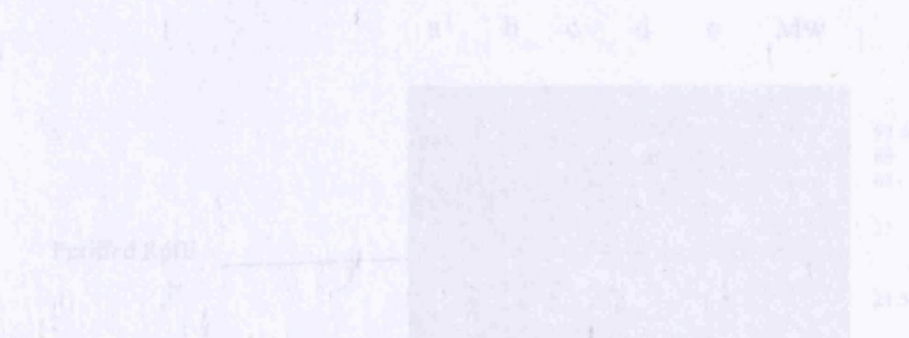


### 3.3.4 Expression and purification of RpfD

This protein had also been found to have a very low level of expression and so as for RpfC, two 150 ml cultures were prepared and incubation continued for 6 h after IPTG induction. A 3  $\mu$ M concentration (300 $\mu$ g/ml) of purified protein was obtained. The gel image is shown in Figure 3.6. The protein band migrated to a position just below the 21.5 kDa markers. The predicted molecular weight is 15,664 Da.

**Figure 3.6**

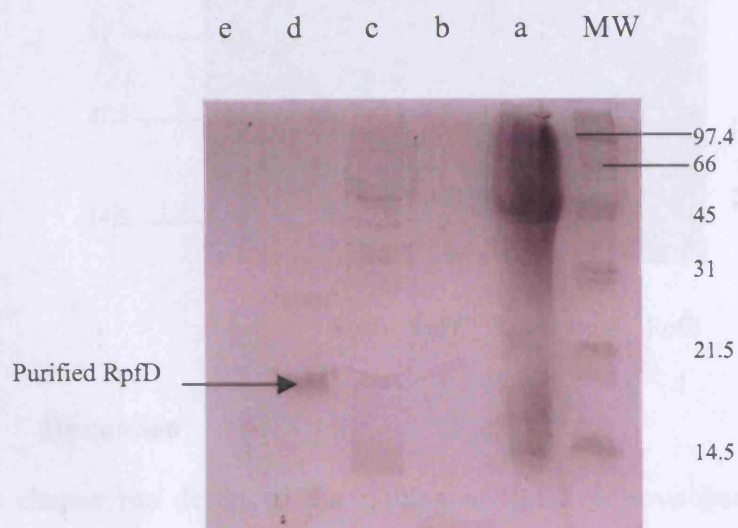
SDS-PAGE showing purified recombinant *M. tuberculosis* RpfD: **a** crude fraction. **b** unbound fraction. **c** wash fraction. **d** elution fraction. **e** strip fraction. A band corresponding to RpfD and molecular weight ~16 kDa is visible in the elution fraction.





**Figure 3.6**

SDS-PAGE Gel showing purified recombinant *M. tuberculosis* RpfD: **a** crude fraction. **b** unbound fraction. **c** wash fraction. **d** elution fraction. **e** strip fraction. A band corresponding to RpfD is visible in the elution fraction.

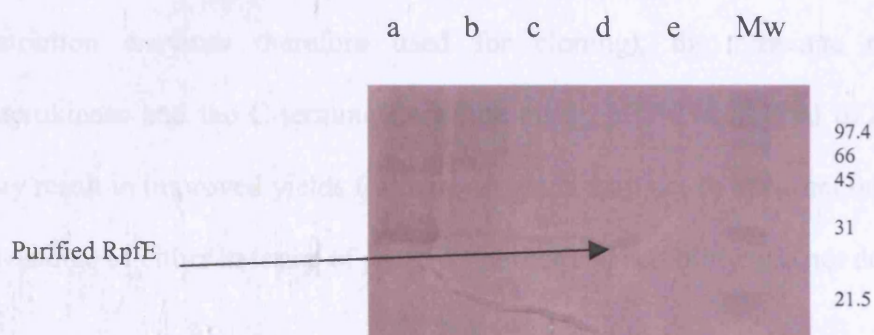


### 3.3.5 Expression and purification of RpfE

Figure 3.7 shows the band representing RpfE in the elution fraction after column purification. The concentration of protein in the elution fraction was calculated as 30  $\mu$ M (560  $\mu$ g/ml). The predicted molecular weight of RpfE is 17,422 Da. On the gel the band appears at just below the 31 kDa marker.

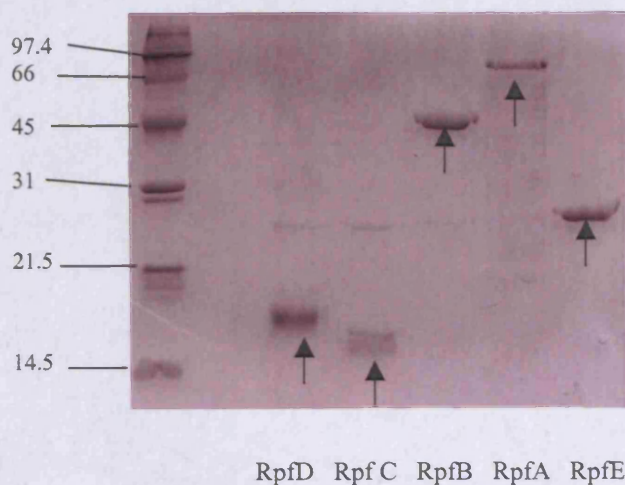
**Figure 3.7**

SDS-PAGE showing purified recombinant *M. tuberculosis* RpfE: **a** crude fraction. **b** unbound fraction. **c** wash fraction. **d** elution fraction. **e** strip fraction. A band corresponding to RpfE with molecular weight ~31 kDa is visible in the elution fraction.



**Figure 3.8**

All five recombinant *M. tuberculosis* Rpfs run on the same gel by SDS-PAGE.



### 3.4 Discussion

This chapter has described the cloning of RpD *de novo* from *M. tuberculosis* DNA, into a host and vector not previously used to express Rpfs (pET-21a and Origami B (DE3) pLysS). It has also described the subcloning of the other four *M. tuberculosis* Rpfs into the Origami B (DE3) pLysS host. As outlined in the introduction, this host, with its *trxB/gor* mutations, permits the formation of disulphide bonds in the cytoplasm and thus may enable optimal folding of a protein that is normally secreted. This host was chosen since it is thought that disulphide bond formation between the cysteine residues found in Rpfs may play an important part in the activity of the proteins. The clones with the Origami B (DE3) pLysS host were used later in the course of the project in an attempt to improve protein solubility and expression (see Chapter 4). There was no particular reason favouring the use of pET-21a rather than pET-19b for cloning RpD. The two vectors differ in various ways including the fusion cloning sites (different restriction enzymes therefore used for cloning), the cleavage of pET-19b by enterokinase and the C-terminal histidine tag in pET-21a. A trial of different vectors may result in improved yields for reasons which may not be apparent but in this case the advantage of either in terms of protein stability and solubility was not demonstrated.

This chapter then described the expression and purification of all five *M. tuberculosis* Rpfs. Concentrations of eluted proteins were low, ranging between 3  $\mu$ M and 30  $\mu$ M, reflecting the poor solubility of these proteins, which will be considered in more detail in the next chapter. The band representing RpfA on SDS-PAGE had a characteristic crenated appearance (see Figure 3.3) which may be associated with the high proline and alanine content of a substantial part of this protein, giving it a snake-like structure. Also it was notable that all five Rpfs, apart from RpfC, when subjected to SDS-PAGE, migrated to a level corresponding to an apparent molecular weight substantially greater than predicted from the sequences – in the case of RpfE the apparent molecular weight was almost twice that predicted. Work published since these experiments were done may shed some light on this (Hartmann 2004). Hartmann's studies of one of the two Rpfs (Rpf2) of *Corynebacterium glutamicum* found that it too had reduced mobility during SDS-PAGE compared with predicted values. Western blot-based glycan detection provided evidence that the Rpf2 was glycosylated. The different mobility seen with different gels might result from differences in the length of the carbohydrate moieties or from the use of different glycosylation sites.

In view of the small concentrations of Rpf eluted after expressing these clones, and the fact that even these were only achieved using 6-8 M urea as a denaturant, the next aim was to improve the solubility of some of the Rpfs, so that they could be produced in larger quantities without the use of a denaturant. Chapter 4 will describe how this was done.

## **CHAPTER IV**

### **RESULTS: IMPROVING SOLUBILITY AND EXPRESSION OF**

**Rpfs**



## 4.1 Introduction

### 4.1.1 Recombinant protein expression in *E. coli*

A wide array of successful protein expression systems based on *E. coli* have been developed, consequent upon our vast knowledge of the genetics and physiology of this organism. However, its use for protein expression does have some drawbacks. For example, it has a limited capacity to secrete proteins, and cannot perform many of the posttranslational modifications found in eukaryotic organisms, such as glycosylation and acetylation. The reducing environment of the cytoplasm creates a problem in terms of the formation of disulphide bonds. Overexpression of heterologous proteins in *E. coli* frequently results in the formation of insoluble aggregated folding intermediates, known as inclusion bodies. Kiefhaber (1991) suggested that the formation of inclusion bodies might be due to the protein translation rate exceeding the capacity of the cell to fold the newly-synthesized polypeptide chain. There may be some advantages to expressing a protein in inclusion body form: large amounts of highly enriched protein, protected from proteolytic degradation, with protection of the host cell from toxic effects of the protein (Clark 2001). However in order to recover protein from such inclusion bodies the covalent bonds that hold the insoluble aggregates together must be disrupted with a denaturant such as 6-8M urea, and this is how the *M. tuberculosis* Rpfs had been produced so far (Chapter 3). The denaturant must then be removed under conditions that are optimal for protein refolding and the final yield of renatured protein may be relatively low. The resulting protein might also be less than optimally active, or not active at all. If Rpfs could be expressed in a soluble form a better yield might be obtained, with improved stability and activity.

Key parameters governing the formation of soluble protein in *E. coli* include structural features of the protein sequence, the stability and translational efficiency of mRNA, the

ease of protein folding, degradation of protein by host cell proteases, major differences in codon usage between the foreign gene and *E. coli*, and plasmid stability (which is reduced if the recombinant protein is toxic to the host) (Old & Primrose 1994, Misawa 1999, Swartz 2001). To help establish a more efficient expression system, two issues are of particular concern – the ‘quantity’ and ‘quality’ of proteins produced. The latter includes consideration of folding, solubility, activity and homogeneity of the protein produced. Factors affecting ‘quantity’ and ‘quality’ of protein are not mutually exclusive. Three strategies were employed to improve protein yield and solubility for the work in this chapter: change of host strain, the use of fusion proteins and lowering the induction temperature.

#### 4.1.2 Host strain

*E. coli* possesses a large number of proteases, located in its cytoplasm, periplasm, inner and outer membranes. Recombinant proteins produced in *E. coli* are particularly prone to proteolysis, as they may appear foreign and are not recognised by the host. *E. coli* can also affect recombinant protein production through rapid degradation of mRNAs, using endonucleases and 3’ exonucleases. Therefore, protease- and endonuclease-deficient host strains such as *E. coli* BL21 (from which Origami B is derived) are commonly used for expressing recombinant proteins. In Chapter 3, Origami B (DE3) pLysS was used to express the RpfS, and this strain is deficient in both *lon* and *ompT* proteases. The product of *lon*, protease La, is one of the most important proteases involved with breakdown of recombinant proteins in *E. coli*. The product of *ompT* is an outer membrane protein, T protease, which cleaves at dibasic amino acid sequences.

The *E. coli* Origami (DE3) hosts have mutations in both thioredoxin reductase (*trxB*) and glutathione reductase (*gor*), permitting the formation of disulphide bonds in the

cytoplasm. In addition the strain used in Chapter 3 had a pLysS plasmid which provides T7 lysozyme to decrease basal expression of target genes and thus stabilise plasmids expressing proteins toxic to *E. coli*.

#### 4.1.3 *Fusion proteins*

By expressing a recombinant protein as a fusion with another protein, better solubility can be achieved. Certain highly soluble fusion protein partners improve the solubility of passenger proteins that would otherwise accumulate as inclusion bodies. Examples of fusion proteins include glutathione S-transferase (GST), thioredoxin (TRX), and NusA (Davis 1999). The fusion protein chosen for use in this work was NusA. This protein has been shown to have good solubilising characteristics and a very high expression level (Davis 1999, Harrison 2000). One of the biological activities of NusA thought to be responsible for its solubilising effect is that of pause-enhancement during transcription (Davis 1999): the presence of NusA causes RNA polymerase to pause for longer periods of time at certain pausing sites along the template DNA. Since transcription and translation are coupled in *E. coli*, NusA may improve solubility by slowing down translation at the transcriptional pauses and allowing critical folding events to occur. The *NusA* gene is available from Novagen as a fusion tag in the vector pET-43.1, under the control of the T7 *lac* promoter. This vector has an N-terminal Nus-tag and an internal histidine tag, with thrombin and enterokinase cleavage sites. A circular map of the plasmid is shown in Figure 4.1.

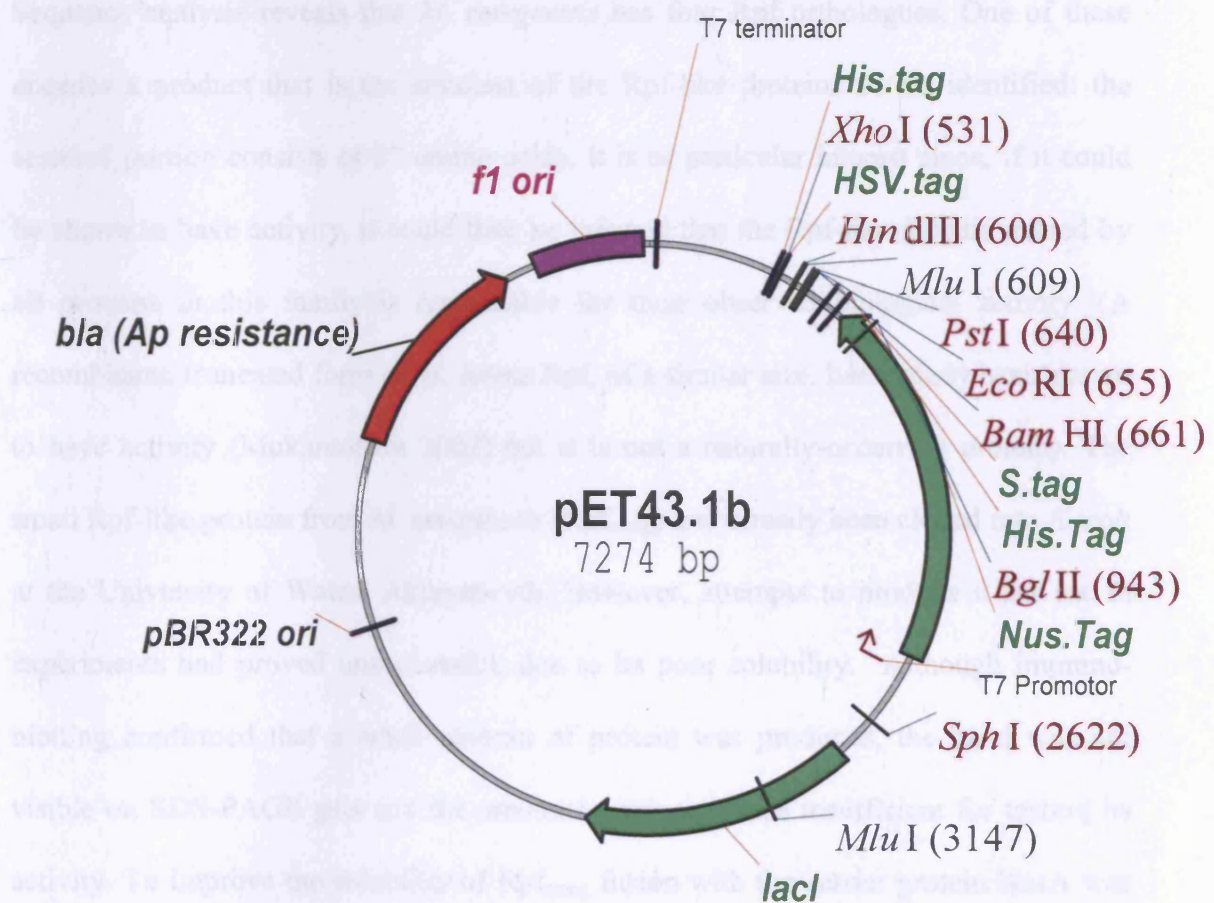
#### 4.1.4 *Induction temperature*

The formation of soluble recombinant proteins in *E. coli* is found to be favoured at growth temperatures lower than 37°C (Schein 1988). This is one of the easiest ways to overcome insoluble protein production. One proposed explanation is that at high

temperatures, heat-shock proteins are produced which include proteases, and this may affect the folding of proteins (Schein 1991). According to Kiefhaber (1991) inclusion body formation is determined by the rate of protein synthesis, folding and aggregation. At higher temperatures, the rate of protein synthesis might be faster, favouring aggregation over folding.

**Figure 4.1**

Diagrammatic representation of vector pET-43.1b (Novagen).



#### 4.1.5 Aims

The aim of the work presented here was to improve the yield and solubility of Rpf. As already mentioned, the strategies adopted involved changes of expression host, the use of a fusion protein, and lowering the induction temperature. Two different sets of

experiments were performed: first, improving the solubility and yield of an Rpf from *Mycobacterium smegmatis* by using a new expression host and a fusion protein (section 4.2) and second, improving the yield and solubility of RpfB and RpfE by changing the expression host and induction temperatures (section 4.3).

## **4.2 Improving the solubility and yield of Rpf from *Mycobacterium smegmatis* by using a new expression host and a fusion protein.**

### **4.2.1 *M. smegmatis* Rpf**

Sequence analysis reveals that *M. smegmatis* has four Rpf orthologues. One of these encodes a product that is the smallest of the Rpf-like proteins so far identified: the secreted portion consists of 87 amino acids. It is of particular interest since, if it could be shown to have activity, it could then be inferred that the Rpf-like domain shared by all proteins in this family is responsible for their observed biological activity. (A recombinant, truncated form of *M. luteus* Rpf, of a similar size, has already been shown to have activity (Mukamolova 2002) but it is not a naturally-occurring protein). The small Rpf-like protein from *M. smegmatis* (Rpf<sub>smeg</sub>) had already been cloned into *E. coli* at the University of Wales, Aberystwyth. However, attempts to produce it for use in experiments had proved unsuccessful, due to its poor solubility. Although immunoblotting confirmed that a small amount of protein was produced, the band was not visible on SDS-PAGE gels and the amounts produced were insufficient for testing its activity. To improve the solubility of Rpf<sub>smeg</sub> fusion with the carrier protein NusA was used. By expressing Rpf<sub>smeg</sub> using a new vector (pET-43.1b - Novagen) and host (Origami (DE3) *E. coli* cells - Novagen) it was hoped that quantities sufficient to demonstrate activity would be obtained. It might also provide a way of producing more stable Rpf, with more reliable expression, as it was hoped that the improved solubility would mean it could be extracted and purified without the use of a denaturant. The host

strain chosen to carry the vector for the work in this chapter was *E. coli* Origami (DE3), from Novagen. Important features of this expression host have already been discussed in 3.1.2 and 4.1.2.

#### 4.2.2 Materials and Methods

Further technical details are to be found in Chapter 2 Materials and Methods 2.1, 2.2 and 2.3 but the following is a summary of the methods used.

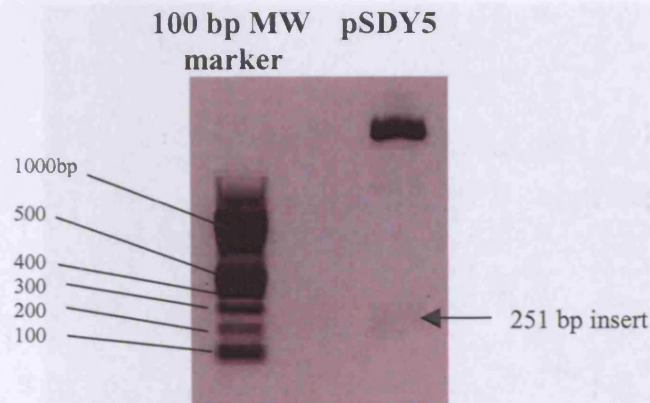
##### 4.2.2.1 Subcloning of *Rpf<sub>smeg</sub>* into *NusA* fusion protein and *Origami* (DE3) host

Clone *E. coli* CA1261 containing plasmid pSDY5 was provided by Danielle Young (University of Wales, Aberystwyth). This plasmid contained a 251 bp fragment encoding the secreted portion of the *Rpf<sub>smeg</sub>* protein. This fragment had been amplified from *M. smegmatis* DNA by PCR with primers 'Smeg 5'(5'-ACC GCG AAT TCG GAC AGC GTG AAC TGG GA-3') and 'Smeg 6'(5'-CAG CTT CTC GAG GCC GCG CTT GCC GCA-3'). The ends of the PCR product had been cut with *Eco*RI and *Xho*I (sites underlined) and cloned into pMTL21 (obtained from N.P.Minton – see Chambers *et al* 1988), cut with the same enzymes. This had been transformed into *E. coli* strain XL2-Blue (Stratagene).

DNA from an overnight culture of this strain was extracted and digested with *Eco*RI and *Xho*I. Agarose gel electrophoresis confirmed the presence of an insert of approximately 250 bp (Figure 4.2.1).

**Figure 4.2.1**

2% agarose gel electrophoresis of plasmid pSDY5 digested with *Eco*RI and *Xho*I to check for the presence of the 251 bp insert encoding *rpf*<sub>smeg</sub>. This is visible as a diffuse band at the expected position.



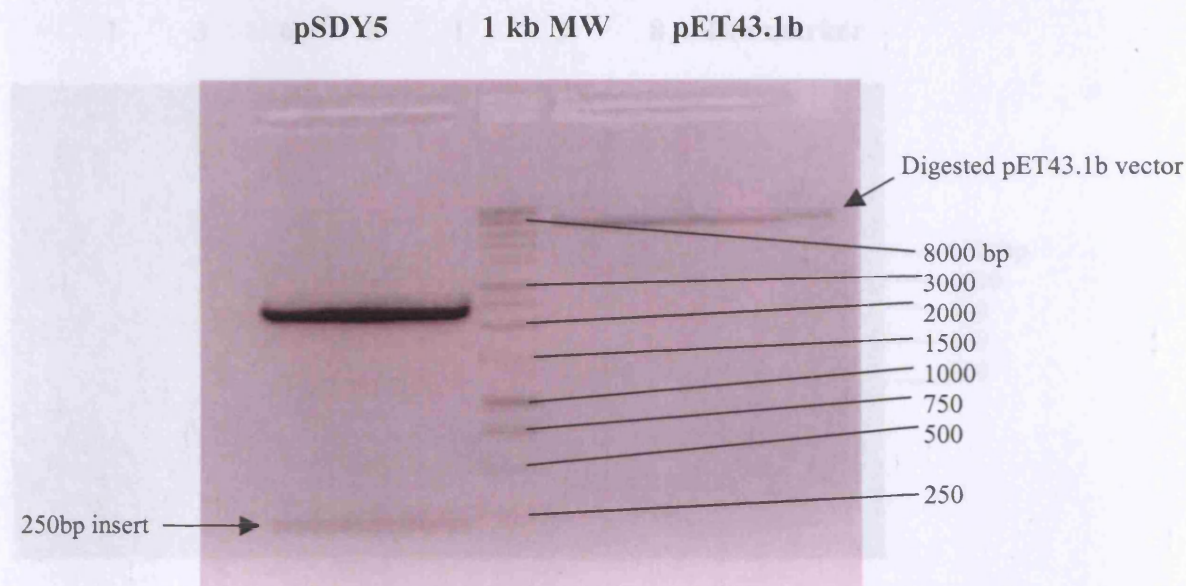
Danielle Young also provided *E. coli* strain CA1272, which comprised *E. coli* strain XL2-Blue (Stratagene) containing the pET-43.1b vector (Novagen). Extracted plasmid DNA was cut within the multiple cloning site with *Eco*RI and *Xho*I, ready for insertion of the pSDY5-derived DNA fragment, encoding the *rpf*-like gene from *M. smegmatis*.

Both the digested vector and the digested plasmid pSDY5 were subjected to electrophoresis. The required insert from pSDY5 was again visible with a size of about 250 bp (Figure 4.2.2). The digested pET-43.1b DNA and the insert from pSDY5 were excised and extracted from the gel. The pET-43.1b vector and the pSDY5 insert were then ligated.



**Figure 4.2.2**

Products of digestion of pSDY5 and pET-43.1b with *Eco*RI and *Xho*I, after extraction from an agarose gel. The 251 bp insert from pSDY5 is visible. 1kb MW marker.

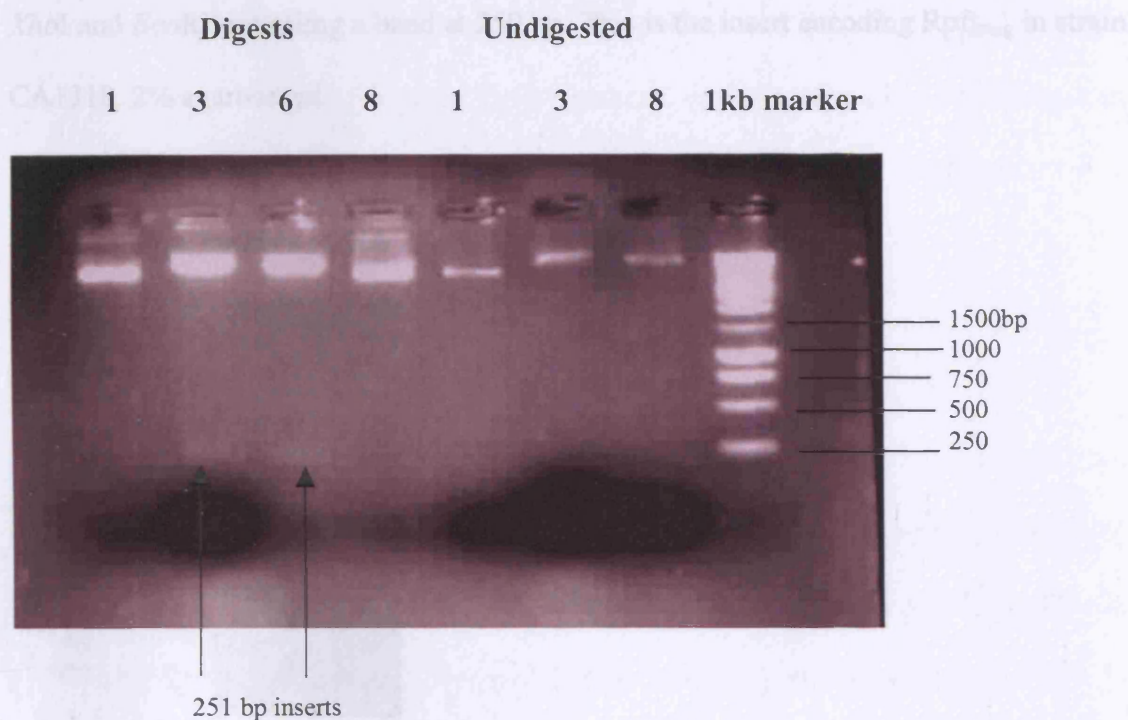


Competent cells of *E. coli* XL2-Blue (Stratagene) were transformed using the ligated DNA (see 2.1.12). Blue-white selection was carried out as described in 2.1.13. The plates yielded white and blue colonies in approximately a 3:1 ratio. White colonies were picked off and subcultured. Plasmid DNA was extracted from four of the transformed *E. coli* clones and digested with *Eco*RI and *Xho*I. The digestion products were separated on an agarose gel. The insert of approximately 250 bp was clearly visible in the clones numbered 3 and 6 in Figure 4.2.3.



**Figure 4.2.3**

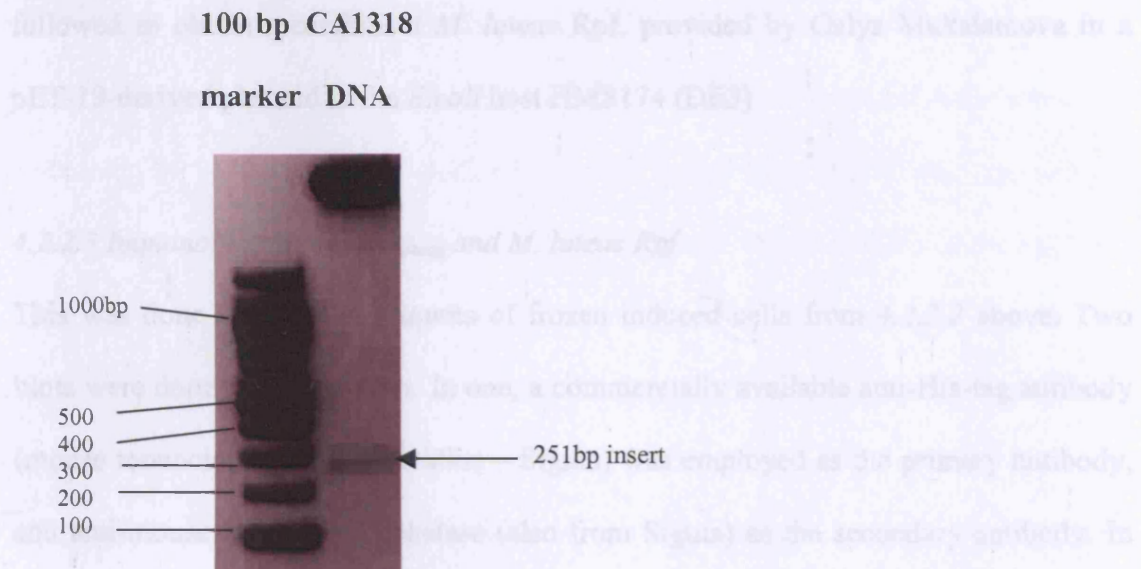
Digestion products of putative clones, containing *rpf<sub>smeg</sub>*: insert visible in clones 3 and 6



DNA extracted from clones 3 and 6 was used to transform competent Origami (DE3) *E. coli* cells (see 2.1.11, 2.1.12). To check that the transformation had been successful, DNA was extracted from four subcultured Origami (DE3) transformants, and digested with *Xho*I and *Eco*RI. Gel electrophoresis revealed bands at 250 bp (Figure 4.2.4) and at 7275 bp, corresponding to the pET-43.1b plasmid (Figure 4.2.5). The new *E. coli* Origami (DE3) strain, harbouring a pET-43.1b derivative containing the *rpf<sub>smeg</sub>* gene, was designated CA1318.

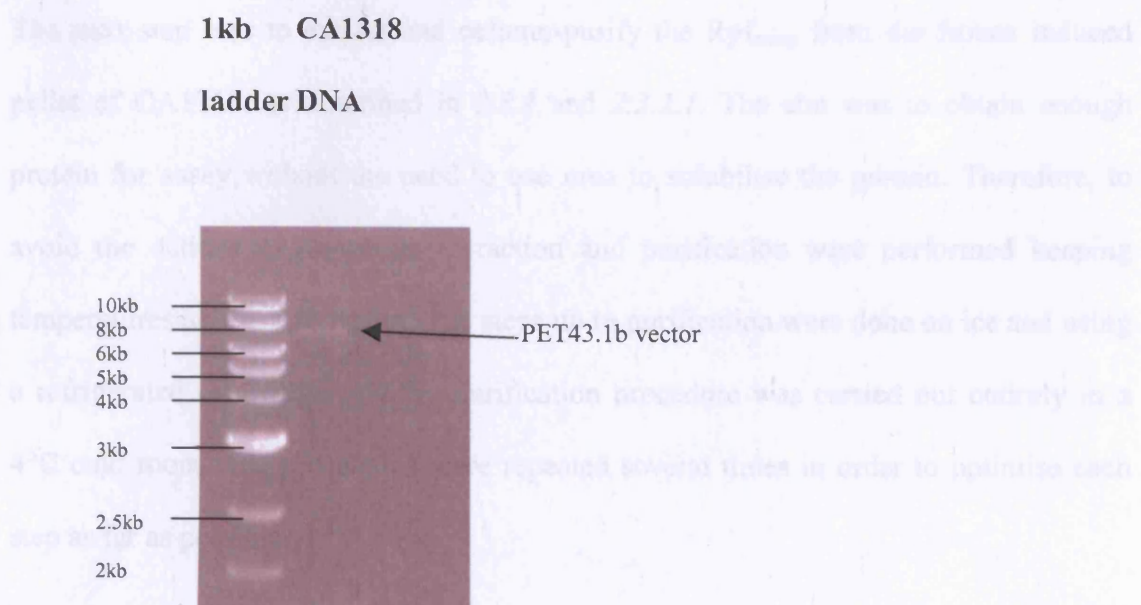
**Figure 4.2.4** Digestion of *Rpf<sub>smeg</sub>* with *M. luteus* Rpf for comparison

DNA extracted from one of the subcultured Origami (DE3) transformants, digested with *Xho*I and *Eco*RI, revealing a band at 250 bp. This is the insert encoding *Rpf<sub>smeg</sub>* in strain CA1318. 2% agarose gel.



**Figure 4.2.5** Digestion of *Rpf<sub>smeg</sub>* with *M. luteus* Rpf for comparison

DNA extracted from the same Origami (DE3) transformant as in Figure 4.2.4, also digested with *Xho*I and *Eco*RI, but run on a 0.8% agarose gel to demonstrate the 7275 bp pET43.1b plasmid.



#### 4.2.2.2 Expression of *Rpf<sub>smeg</sub>* with *M. luteus* Rpf for comparison

150 ml cultures of CA1318 (containing the plasmid-encoded *M. smegmatis* *rpf* gene), were induced with IPTG, then harvested and frozen (2.2.1, 2.2.2). For later SDS-PAGE comparison with an Rpf of known yield and band mobility, the same procedure was followed to obtain recombinant *M. luteus* Rpf, provided by Galya Mukalamova in a pET-19-derived plasmid in the *E.coli* host HMS174 (DE3).

#### 4.2.2.3 Immunoblotting for *Rpf<sub>smeg</sub>* and *M. luteus* Rpf

This was done using 20 µl aliquots of frozen induced cells from 4.2.2.2 above. Two blots were done simultaneously. In one, a commercially available anti-His-tag antibody (mouse monoclonal antipolyhistidine – Sigma) was employed as the primary antibody, and anti-mouse alkaline phosphatase (also from Sigma) as the secondary antibody. In the other, a primary specific anti-Rpf antibody was employed (obtained from Mike Young's collaborators in Moscow). This mouse antibody detects the conserved Rpf-like region of the protein. The secondary antibody used was the same.

#### 4.2.2.4 Extraction and purification of *Rpf<sub>smeg</sub>*

The next step was to extract and column-purify the *Rpf<sub>smeg</sub>* from the frozen induced pellet of CA1318, as described in 2.3.1 and 2.3.2.1. The aim was to obtain enough protein for assay without the need to use urea to solubilise the protein. Therefore, to avoid the actions of proteases, extraction and purification were performed keeping temperatures at 4°C throughout. The steps up to purification were done on ice and using a refrigerated centrifuge, and the purification procedure was carried out entirely in a 4°C cold room. The procedures were repeated several times in order to optimise each step as far as possible.



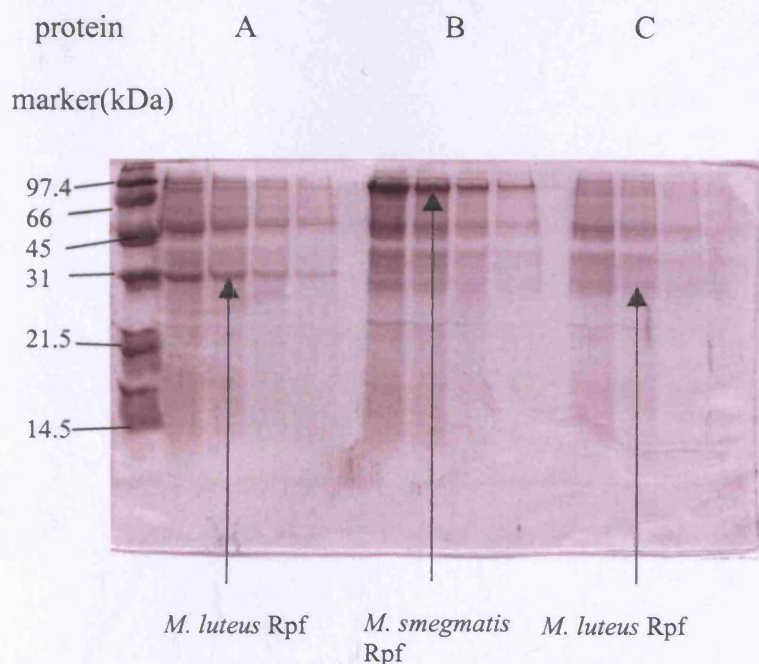
### 4.2.3 Results

#### 4.2.3.1 Expression of *Rpf<sub>smeg</sub>*

To check that CA1318 was producing *Rpf<sub>smeg</sub>*, a small sample of the frozen induced cells from 4.2.2.2 was taken and diluted in sample buffer. It was boiled, centrifuged at 13,000 rpm for 10 min and then the supernatant was loaded on a SDS-PAGE gel at various dilutions (undiluted, 1:2, 1:5 and 1:10). This was repeated with the clone containing *M. luteus* Rpf. After SDS-PAGE, clear bands were visible, corresponding in size those expected for the histidine-tagged recombinant versions of *M. luteus* Rpf (~31 kDa) and *Rpf<sub>smeg</sub>* (~70 kDa) (Figure 4.2.6).

**Figure 4.2.6**

SDS-PAGE gel showing crude extracts of: A-Recombinant *M. luteus* Rpf (pET-19 vector, HMS174 (DE3) host); B-Recombinant *Rpf<sub>smeg</sub>* CA1318 (pET-43.1b vector, Origami (DE3) host); C-Recombinant *M. luteus* Rpf (pET-19 vector, Origami (DE3) host). All are shown in dilutions of, from left to right, neat, 1:2, 1:5, 1:10.

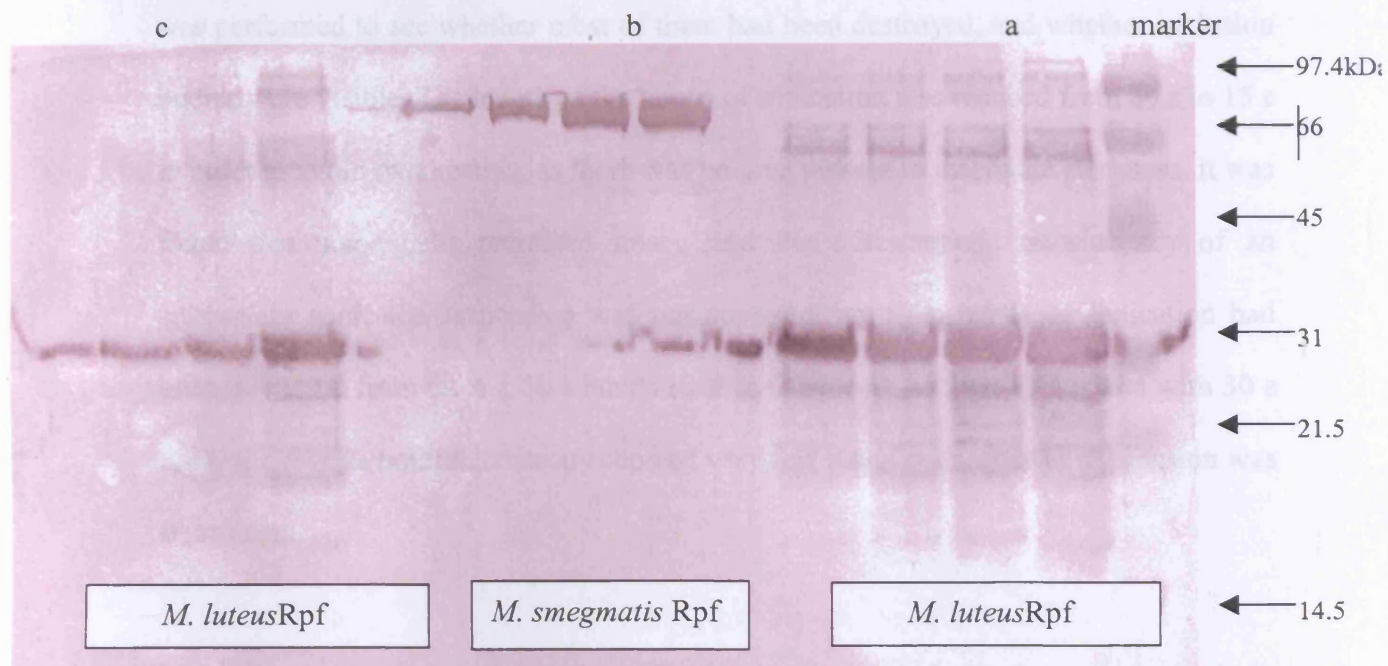


To confirm that this was the desired Rpf protein, immunoblotting was carried out. The results are shown in Figure 4.2.7. Using both the anti-His-tag and the anti-Rpf antibodies, bands representing Rpf<sub>smeg</sub> and *M. luteus* Rpf were clearly visible. Using the anti-His-tag antibody, but not the anti-Rpf antibody, multiple bands were seen with the Rpf<sub>smeg</sub>. These additional bands might represent C-terminal degradation products, which would contain a His-tag but not Rpf (hence they were not visible using the anti-Rpf antibody). With the Rpf-specific blot, numerous bands were visible with the *M. luteus* Rpf. As the anti-Rpf serum has not been affinity purified, it may be less pure and specific than the anti-His-tag antibody.

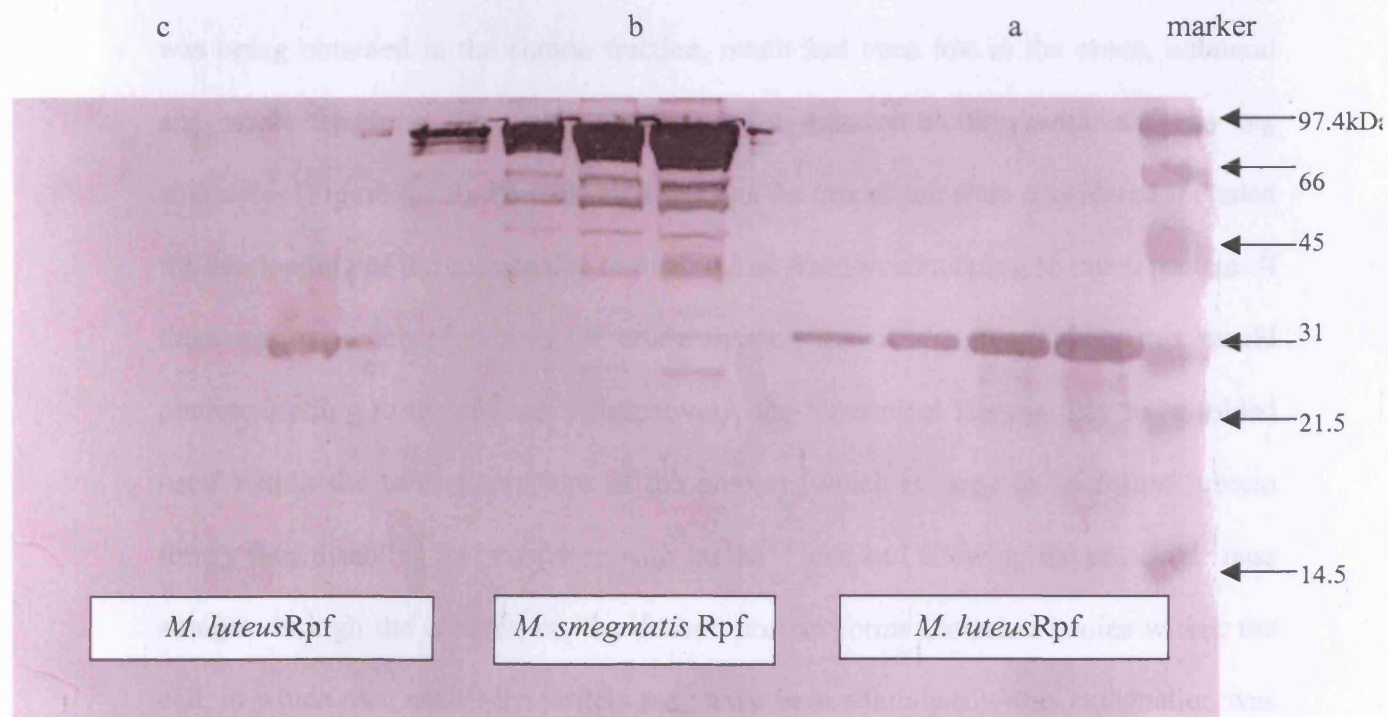
**Figure 4.2.7**

Immuno-blot of the SDS-PAGE gel in Figure 4.2.6

*A* stained with anti-Rpf antibody



*B* stained with anti-His-tag antibody



#### 4.2.3.2 Extraction and purification of *Rpf<sub>smeg</sub>*

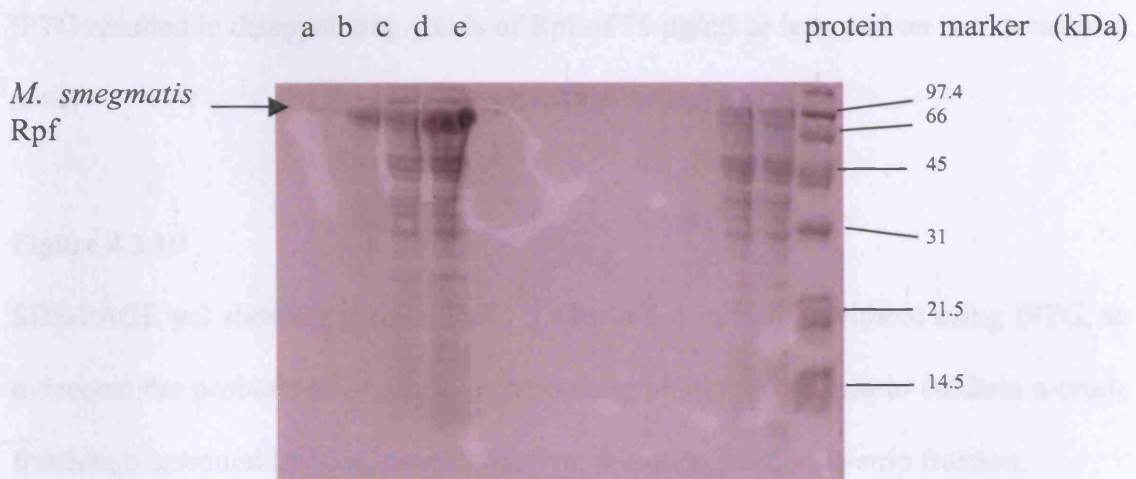
This was performed seven times in order to establish the optimum conditions for the greatest yield and solubility. Each time, after sonication on ice, microscopy of the cells was performed to see whether most of them had been destroyed, and whether inclusion bodies were visible. The length of the bursts of sonication was reduced from 30 s to 15 s in order to avoid overheating, as there was no urea present to inactivate proteases. It was found that some cells remained intact, and the characteristic translucency of an adequately sonicated suspension was not apparent, until the bursts of sonication had been increased from the 6 x 30 s bursts used in Chapter 3, to 20 x 15 s bursts with 30 s intervals. At this point microscopy showed very few intact cells, and the suspension was translucent.

Initial attempts at expression and purification resulted in a large amount of protein in the crude extract, as seen on SDS-PAGE in Figure 4.2.8. Although ~100 µg/ml pure Rpf was being obtained in the elution fraction, much had been lost in the crude, unbound and wash fractions. This was confirmed by immuno-blotting with anti-His tag antibodies (Figure 4.2.9). Possible explanations for this which were considered included the overloading of the column due to the original fraction containing so much protein. If there was so much protein in the crude fraction that it was aggregating this would prevent binding to the column. Alternatively, the N-terminal His-tag may have folded itself within the tertiary structure of the protein (which is large in its fusion protein form), thus disabling its interaction with the Ni<sup>2+</sup> ions and allowing the protein to pass straight through the column; or, the desired protein forms inclusion bodies within the cell, in which case not all the protein may have been solubilised – this explanation was not borne out by microscopy of the sonicated cells, which did not appear to contain inclusion bodies.



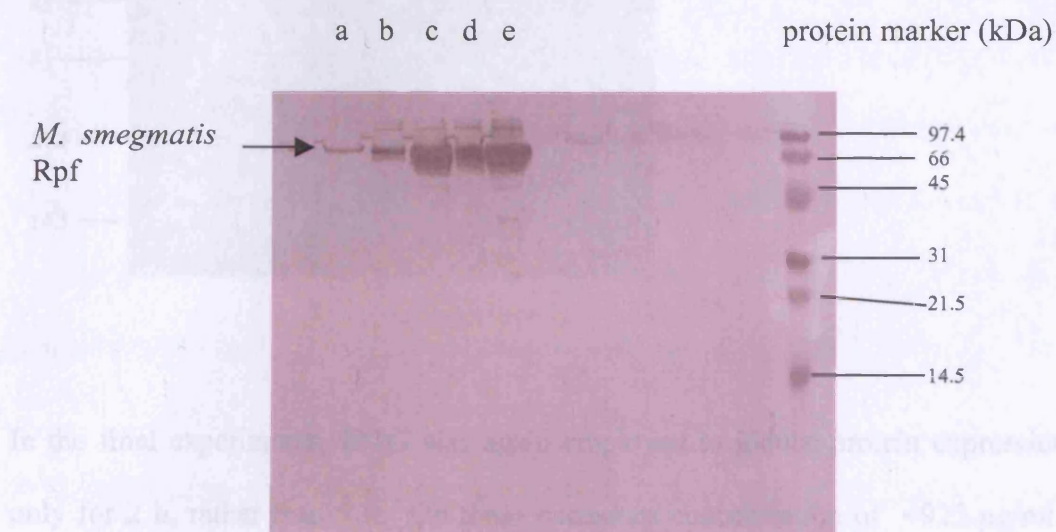
**Figure 4.2.8**

SDS-PAGE gel of CA1318 protein extract showing recombinant Rpf<sub>smeg</sub> before optimisation of expression: aggregation of protein in the crude fraction may have prevented binding to the column. a-strip fraction; b-elution fraction; c-wash fraction; d-unbound fraction; e-crude fraction.



**Figure 4.2.9**

Immunoblot of gel in Figure 4.2.8 above, showing most Rpf<sub>smeg</sub> is lost in the crude fraction and a smaller amount remains in the elution fraction. a-strip fraction; b-elution fraction; c-wash fraction; d-unbound fraction; e-crude fraction.



Considering the first possibility - that there might be so much protein in the crude fraction that it was aggregating and not binding to the column - further expression

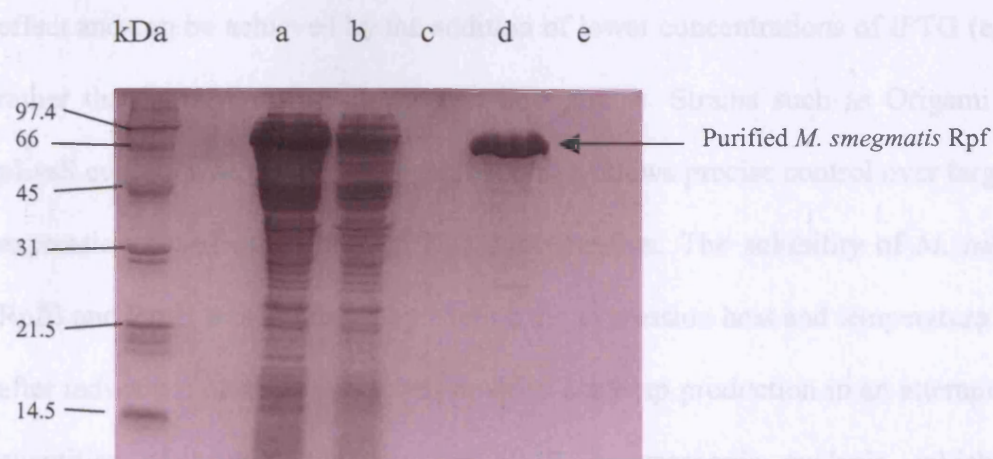


experiments were performed with a five hour culture time but without the addition of IPTG (which induces high-level expression). On one occasion a protein concentration of 235  $\mu\text{g/ml}$  was obtained in the elution fraction and a large band was visible in this fraction (Figure 4.2.10). However, a relatively large amount of the protein was still lost in the unbound and wash fractions, and on three other occasions expression without IPTG resulted in disappointing yields of Rpf of 75  $\mu\text{g/ml}$  or less, and on two occasions, none.

**Figure 4.2.10**

SDS-PAGE gel showing purified Rpf<sub>smeg</sub> after a 5 h induction without using IPTG, to overcome the problem of aggregation preventing binding of protein to column. a-crude fraction; b-unbound fraction; c-wash fraction; d-elution fraction; e-strip fraction.

protein marker



In the final experiments, IPTG was again employed to induce protein expression, but only for 2 h, rather than 3 h. On these occasions concentration of  $\sim 925 \mu\text{g/ml}$  were obtained, with a strong band visible on SDS-PAGE. Thus the optimum procedure for obtaining a good yield of protein in the elution fraction, in the absence of a denaturant,

was found to involve the addition of IPTG for protein expression, with incubation for 2 h afterwards, with sonication for 8 or more bursts of 15 s, as necessary, with 30 s intervals to prevent heating. The suspension should be translucent, and microscopy should demonstrate that cells have been completely disrupted.

This was the first time Rpf from *Mycobacterium smegmatis* was expressed in quantities adequate for use in assays, achieved by using a suitable expression host and a vector which created a soluble fusion protein.

#### **4.3 Improving the yield and solubility of RpfB and RpfE by changing the expression host and induction temperatures, and scaling up of production to achieve greater yields.**

As already discussed, lowering the temperature of growth increases the yield of correctly folded, soluble protein; a slower induction of protein expression has a similar effect and can be achieved by the addition of lower concentrations of IPTG (eg 0.1 mM rather than 1 mM) in certain *E. coli* host strains. Strains such as Origami B (DE3) pLysS contain a *lacY* deletion mutation which allows precise control over target protein expression based on inducer (IPTG) concentration. The solubility of *M. tuberculosis* RpfB and RpfE was improved by altering the expression host and temperature of culture after induction; attempts were then made to scale-up production in an attempt to obtain quantities of protein adequate for NMR spectroscopic analysis, which requires milligram quantities of protein - far in excess of those achieved in Chapter 3.

RpfB was chosen for these experiments as it was easier to produce in more substantial amounts, at least in the presence of a denaturant; RpfE was chosen because of its

smaller size (molecular weight 17.4 kDa) which made it a more suitable potential subject for NMR.

#### *4.3.1 Materials and Methods*

Further technical details are to be found in Chapter 2 Materials and Methods **2.1**, **2.2** and **2.3** but the following is a summary of the methods used.

##### *4.3.1.1 Generating clones of *rpfB* and *rpfE* in a new host and vector to improve solubility*

*rpfB* and *rpfE* cloned into plasmid vector pET-19b were used in a new host, Origami B (DE3) pLysS (Novagen) (see 3.2.2). The features of this expression host are outlined in 3.1.2. A further clone was constructed by subcloning *rpfE* from pET-19b into pET-15b.tev using restriction enzymes *Bam*HI and *Nde*I. The pET-15b.tev vector allows removal of the N-terminal His-tag using TEV protease. Again Origami B (DE3) pLysS was used as the host.

*rpfB* in pET-19b vector and Origami B (DE3) pLysS host was designated clone AD5; *rpfE* in pET-19b vector and Origami B (DE3) pLysS host was designated clone AD8; *rpfE* in pET-15b.tev vector and Origami B (DE3) pLysS host was designated clone AD11. A table of clones used in this project is shown in Appendix *iii*.

##### *4.3.1.1 Expression of *RpfB* and *RpfE* from clones AD5, AD8 and AD11*

150ml cultures of the clones were induced with 0.1 mM IPTG (as opposed to 1 mM used in Chapter 3) and expressed at temperatures from 18°C - 22°C, for 16 h (unless stated otherwise). During extraction, none of the buffers contained urea, and sonication consisted of 15 s bursts with 30 s intervals, rather than the 30 s bursts used when urea

was present in the buffers, to minimise overheating. To minimise degradation by proteases, all buffers were ice-cold, centrifugation was done in a refrigerated centrifuge at 4°C, and the protein samples were kept on ice at all stages. The protein was extracted in this way on several occasions. In some experiments a protease inhibitor cocktail (Roche) was used.

#### *4.3.1.2 Large scale expression of RpfE from clone AD11*

It was decided to scale up the expression from clone AD11 (*rpfE* in pET-15b.tev vector and Origami B (DE3) pLysS host). This part of the work was carried out at Birkbeck College with the co-operation of Andrei Okorokov. The procedure is described in full in Chapter 2 Materials and Methods section 2.3, specifically 2.3.4 and 2.3.5. Eight litres of culture of the *E. coli* host were set up for induction and IPTG 0.1 mM was used. The flasks were then incubated at 25°C for a further 4 ½ hours. Protease Inhibitor Cocktail (Roche) was added. The cell extract was loaded onto an AKTA™ FPLC™ chromatography system (Amersham Biosciences) and run down two 5 ml HiTrap nickel-chelating HP columns™ (Amersham Biosciences). Desalting was performed on the same equipment using HiPrep 26/10 50ml desalting columns™ (Amersham Biosciences) and binding buffer. The use of pET-15b.tev meant that eluted protein could be treated with TEV protease to remove the His-tag from the RpfE (2.3.5).

#### *4.3.2 Results*

##### *4.3.2.1 Expression of RpfB from clone AD5 (rpfB in pET-19b vector and Origami B (DE3) pLysS host)*

150 ml cultures of clone AD5 induced with 0.1 mM IPTG were expressed at room temperature (22°C) for 16 hours, and extracted at low temperatures as described in section 4.3.1.2. RpfB was extracted in this way on several occasions. The amounts of protein in the eluted fractions ranged from 60 - 125 µg/ml. This compared with yields of

400 - 500 µg/ml when extracted with urea. SDS-PAGE confirmed a protein band at ~45 kDa, corresponding to RpfB. The use of the protease inhibitor cocktail (Roche) had no effect on the final protein yield (data not shown).

#### *4.3.2.2 Expression of RpfE from clone AD8 (rpfE in pET-19b vector and Origami B (DE3) pLysS host)*

150 ml cultures of clone AD8 were induced with 0.1 mM IPTG and expressed at room temperature (22°C) for 16 hours. Extraction was performed as described for clone AD5. This was done twice and the amount of protein obtained in the elute fractions was 50 µg/ml and 60 µg/ml. SDS-PAGE confirmed a protein band at ~31 kDa, corresponding to RpfE.

#### *4.3.2.3 Expression of RpfE from clone AD11 (rpfE in pET-15b.tev vector and Origami B (DE3) pLysS host)*

A 150 ml culture of AD11 was induced with 0.1 mM IPTG at room temperature for 16 hours and extracted as described above. 50 µg/ml of recombinant protein was obtained in the eluted fraction and SDS-PAGE confirmed a faint band at ~31 kDa, again corresponding to RpfE. Therefore there appeared to be no difference in yield between clones AD8 and AD11, with the two different vectors pET-15b.tev and pET-19b. This experiment was repeated using a 500ml culture of clone AD11. Again 0.1 mM IPTG was used, but this time induction was for 5h at 24°C. Protease inhibitors were added to all the buffers, and prior to sonication lysozyme was added at a concentration of 200 µg/ml, to improve the extraction. In this experiment the eluted fraction contained a recombinant protein concentration of 100 µg/ml, which was disappointing and represented an even poorer yield even than from the 150ml culture.

These results demonstrated that soluble protein could be obtained from these clones - through the use of the Origami B (DE3) pLysS host, lower induction temperatures and lower concentrations of inducers - albeit in amounts approximately ten times less than was achieved using urea. Clones AD8 and AD11 yielded a similar amount of recombinant RpfE (~50 µg/ml in the eluted fractions, or approximately 150 µg from a 150 ml culture) despite making use of different vectors. Clone AD5 gave a slightly better yield of RpfB – 60-125 µg/ml – but still much less than the 400-500 µg/ml amounts extracted using urea.

#### *4.3.2.4 Scaling-up of production of RpfE*

In order to obtain a larger quantity of protein, eight litres of culture of AD11 were induced. Extraction was done on the deposit from 2 l culture volumes - otherwise the volume of frozen deposit would have been too great for effective cell lysis and protein extraction. After running the extracted material down the nickel columns, aliquots of the eluate which corresponded to protein peaks on the chromatogram were saved and kept at 4°C. Aliquots containing protein peaks were aliquots numbers 40-45 and 51 –56 out of 66 aliquots.

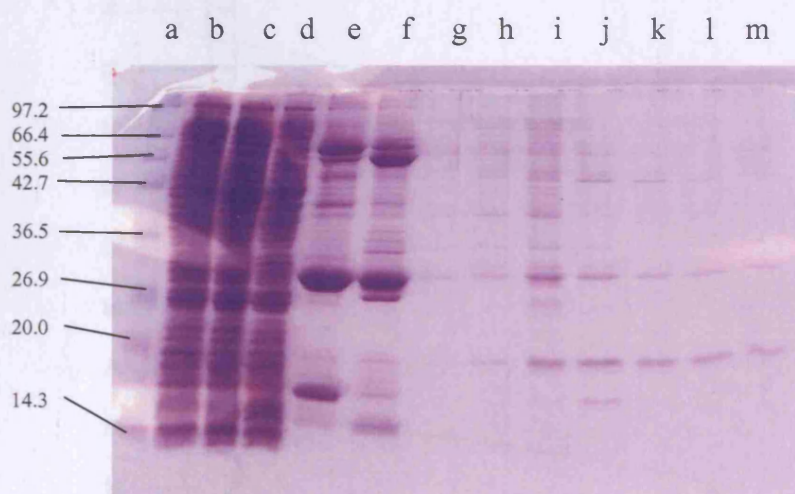
15 µl of each of these aliquots was run on a 15% SDS-PAGE gel. This showed ~ 28 kDa bands and ~19 kDa bands in the aliquots numbered 40-45 and 51-56 (Figure 4.3.1a and b). RpfE on SDS-PAGE gels runs at just below the 31 kDa marker and could correspond to the ~28 kDa band seen. Most of the ~28 kDa protein was seen in fractions 52 and 53; the protein in fractions 40-45 was mainly ~19 kDa – too small to represent RpfE. Therefore it appeared that most of the RpfE was in fractions 51-56. Nonetheless, fractions 40-45 were pooled (pooled volume 40 ml) and fractions 51-56 were pooled (pooled volume 40 ml). Both pools were then desalted. The resulting chromatograms

are shown in Figures 4.3.2 and 4.3.3. The protein flows through the desalting column before the imidazole in the sample, so that they become separated, a process reflected in the  $A_{280}$  of the eluted material which shows two peaks, the first corresponding to the protein and the second to the imidazole. After desalting, the  $A_{280}$  was measured for each protein pool, 1 unit being equivalent to 1 mg/ml protein. The protein content of pooled fractions 40-45 was 3.4 mg, and of fractions 51-56, 2.24 mg.

**Figure 4.3.1a**

SDS-PAGE of aliquots of eluted protein obtained from clone AD11 after passing through the FPLC equipment. ~28 kDa and ~19 kDa bands are present in fractions 40-45.

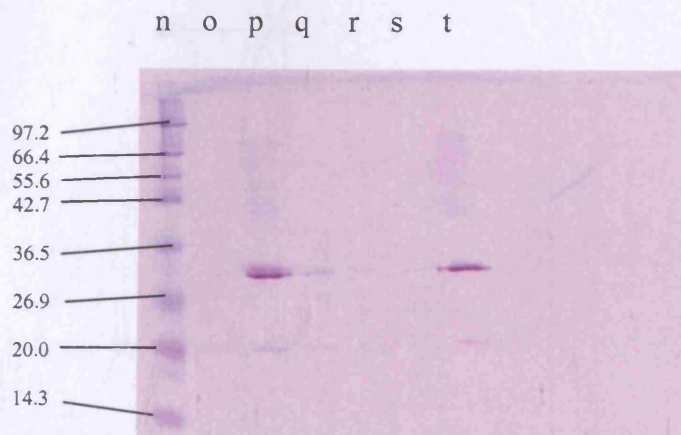
a	molecular weight marker	g	fraction 39
b	sonicated cell deposit (clone AD11)	h	fraction 40
c	crude extract from AD11	i	fraction 41
d	flow-through (unbound)	j	fraction 42
e	fraction 21	k	fraction 43
f	fraction 22	l	fraction 44
		m	fraction 45



**Figure 4.3.1b**

SDS-PAGE of aliquots of eluted protein obtained from clone AD11 after passing through the FPLC equipment. ~28 kDa and ~19 kDa bands are present in fractions 51-56.

n	molecular weight marker	q	fraction 54
o	fraction 46	r	fraction 55
p	fraction 53	s	fraction 56
		t	fraction 52

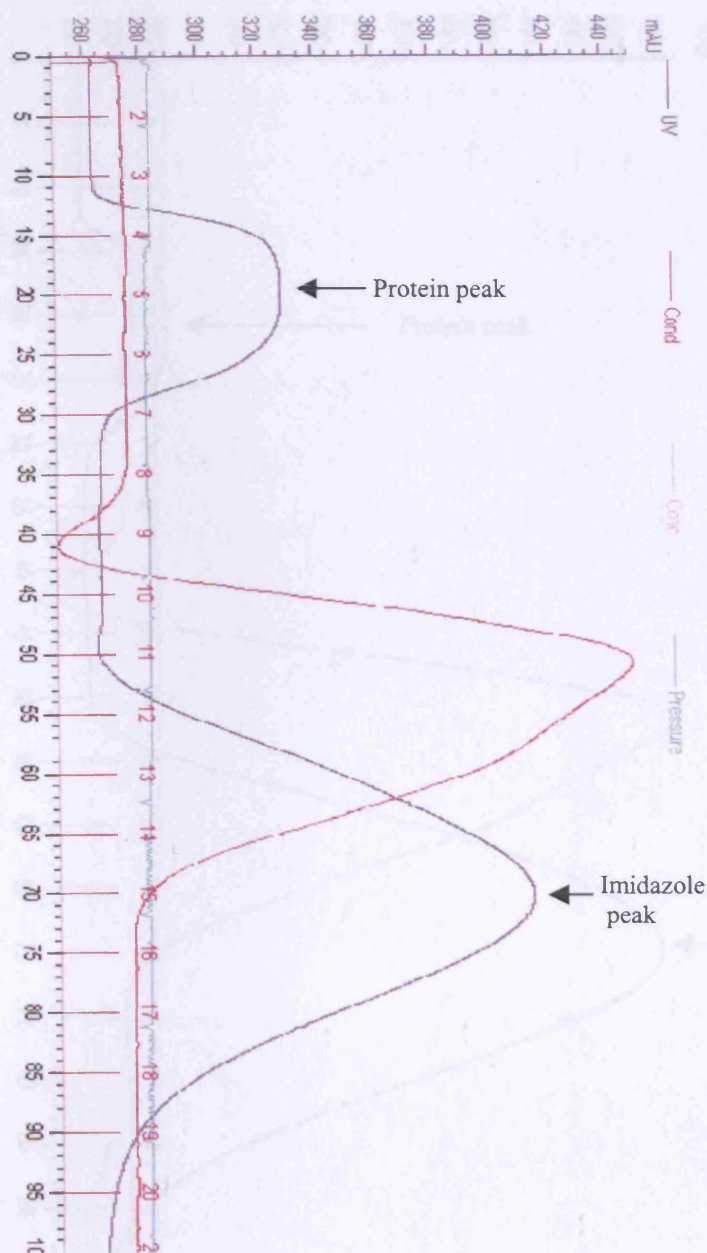




**Figure 4.3.2**

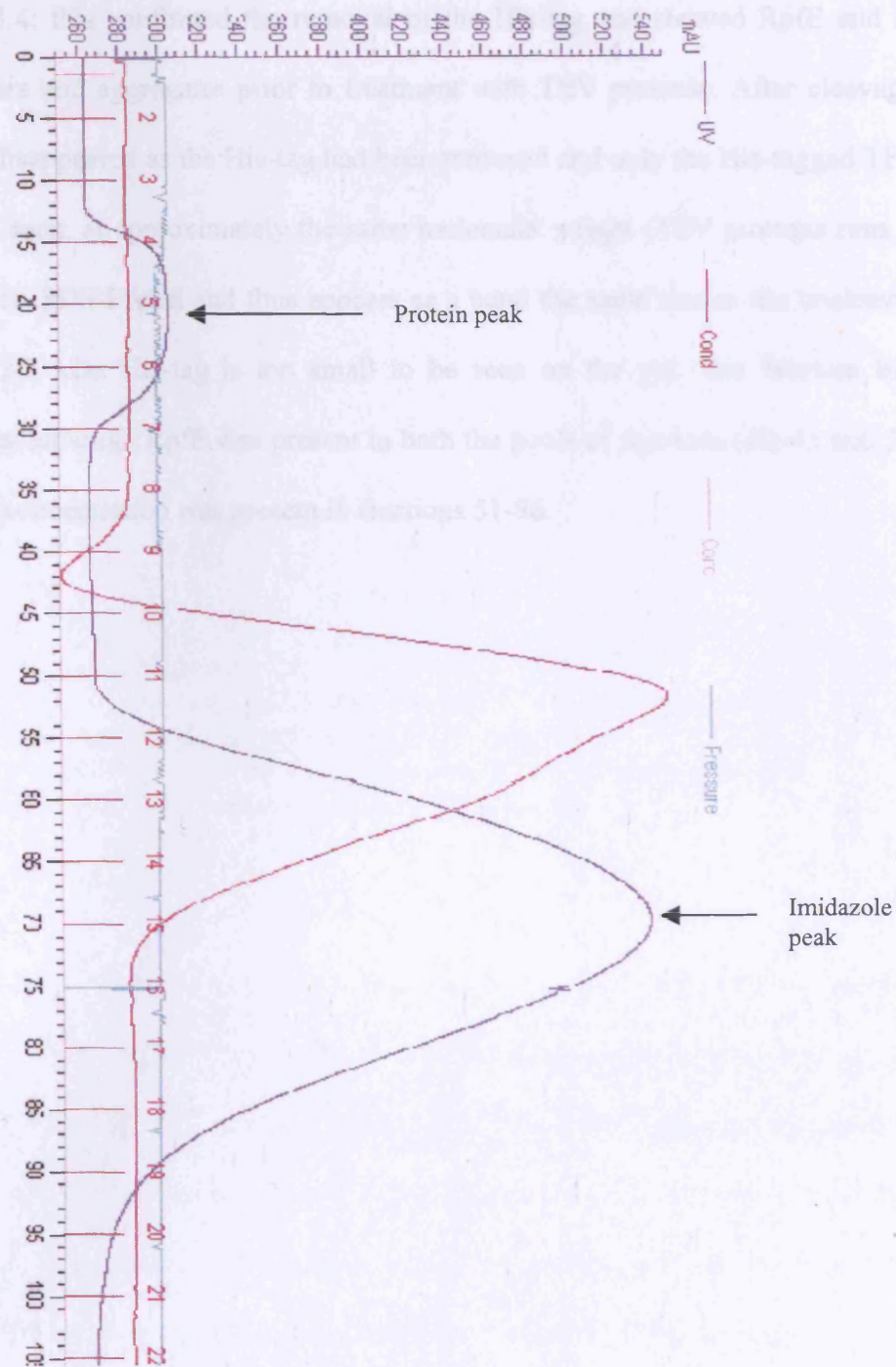
Chromatogram of desalting of pooled protein fractions 40-45 on Amersham AKTA

FPLC equipment. The dark blue line represents UV absorbance and two peaks are seen corresponding to eluted protein and imidazole.



**Figure 4.3.3**

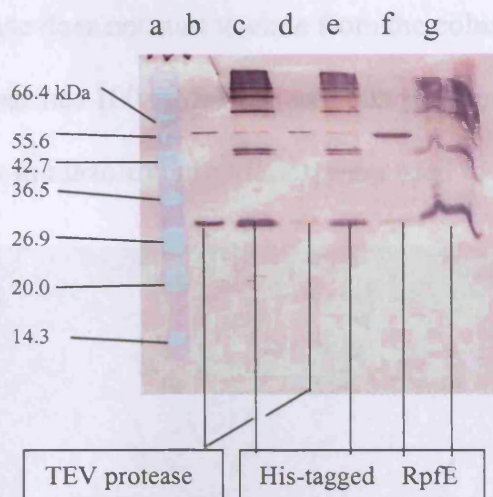
Chromatogram of desalting of pooled protein fractions 51-56 on Amersham AKTA FPLC equipment. The dark blue line represents UV absorbance and two peaks are seen corresponding to eluted protein and imidazole.



Each of the two protein pools were then treated with TEV protease to remove the His-tag from the RpfE (2.3.5). Samples were removed from the reaction hourly for the first 3 hours and then after overnight incubation. SDS-PAGE of these samples showed that the removal of the His-tag was not evident at 3 h after the start of the reaction, but had taken place by 24 h. A Western blot of the reaction, using anti-His antibodies, is shown in Figure 4.3.4: this confirmed the removal of the His-tag and showed RpfE and its dimers, trimers and aggregates prior to treatment with TEV protease. After cleavage, these bands disappeared as the His-tag had been removed and only the His-tagged TEV protease was seen, at approximately the same molecular weight (TEV protease runs to ~26-29 kDa on SDS-PAGE and thus appears as a band the same size as the uncleaved RpfE). The 2-3 kDa His-tag is too small to be seen on the gel. The Western blot confirmed that although RpfE was present in both the pools of fractions (40–45 and 51-56), a higher concentration was present in fractions 51-56.

**Figure 4.3.4**

Immunoblot of fractions 40-45 and 51-56 with and without overnight digest with TEV protease, using anti-His-tag antibodies. Without cleavage, RpfE at ~28 kDa is seen, together with dimers, trimers and aggregates of RpfE. With cleavage, the His-tagged TEV protease is visible with the same molecular weight; the cleaved RpfE is no longer visible as the His-tag has been removed. 2-3 kDa His-tags are too small to be visible on this gel.



Lanes:

- a molecular weight marker
- b pooled desalted fractions 51-56 after overnight TEV protease cleavage
- c pooled desalted fractions 51-56 without TEV protease
- d pooled desalted fractions 40-45 after overnight TEV protease cleavage
- e pooled desalted fractions 40-4 without TEV protease
- f crude cell extract
- g crude cell deposit

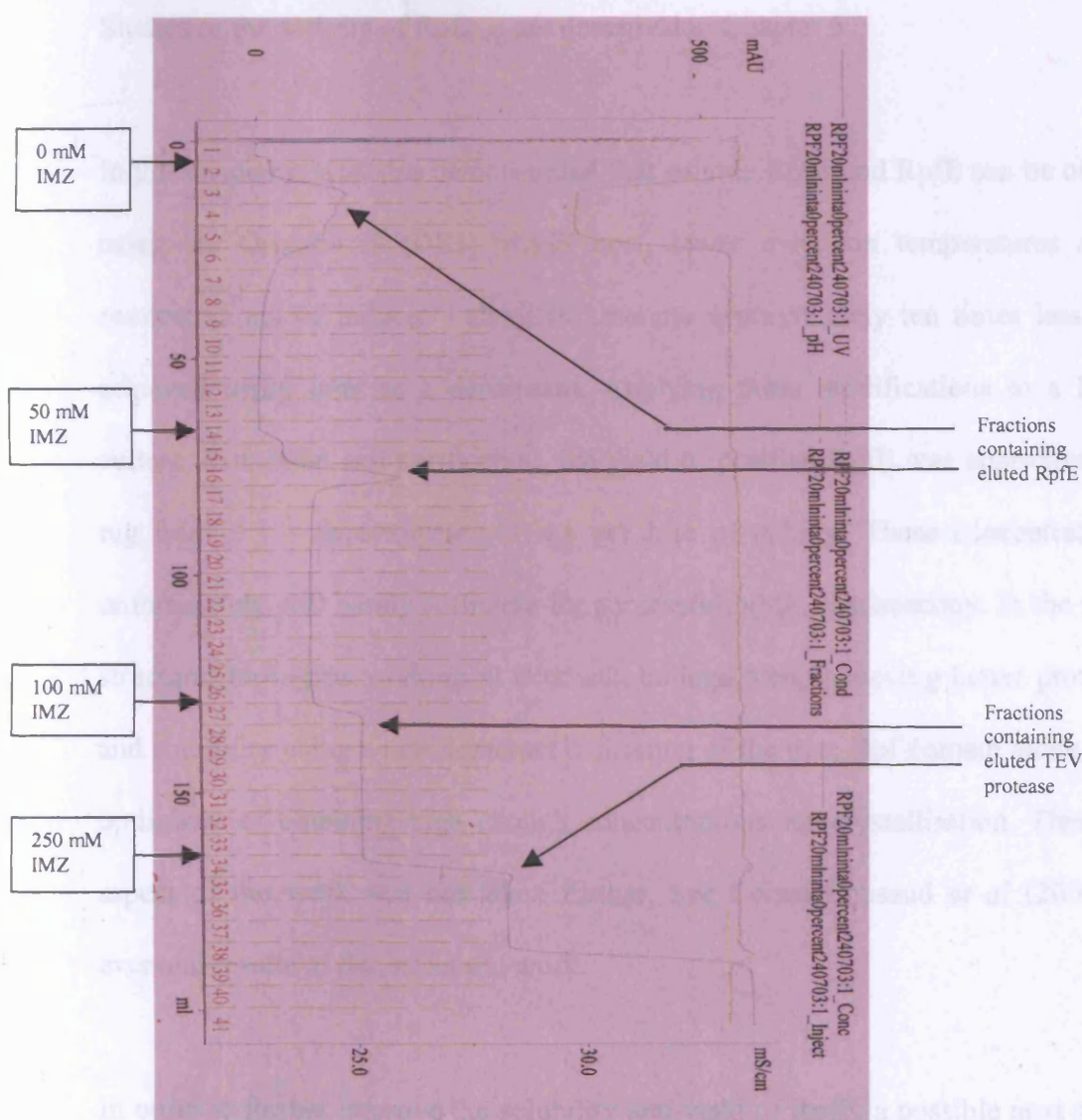
The two pooled protein fractions were then run down the FPLC column again to separate the TEV protease from RpfE: as TEV protease is His-tagged it binds the column while RpfE passes through (Figure 4.3.5). Since Rpf has some non-specific binding, peaks were collected after running through elution buffer containing increasing concentrations of imidazole (the His-tagged TEV protease would not be expected to detach until 100 mM imidazole concentration was reached). The aliquots of eluate corresponding to the peaks occurring below 100 mM imidazole thus contained purified RpfE, extracted without a denaturant, from which the His-tag had been removed.



The overall yield of RpfE obtained from a 2 l culture volume was ~2.25 mg – about 1 mg/l.

**Figure 4.3.5**

Chromatogram showing separation of RpfE from TEV protease using nickel columns on Amersham AKTA FPLC equipment, eluting with imidazole (IMZ). The His-tagged TEV protease does not start to elute from the column until the concentration of imidazole reaches 100 mM; RpfE exhibits some non-specific binding and so step-wise increases in imidazole concentration were used to elute it.



#### 4.4 Discussion

In this chapter attempts were made to improve the solubility and yield of three different Rpf<sub>s</sub>, and to extract them without using a denaturant. By expressing Rpf<sub>smeg</sub> as a fusion protein with NusA, a much greater quantity of soluble protein could be extracted, without the use of urea. Previous attempts to express this protein had resulted in quantities too small to visualise on a Coomassie-Blue stained gel or on immunoblotting. Quantities sufficient for testing of biological activity were now produced: and if found to be active, this would support the belief that the core 'Rpf domain' common to all Rpf<sub>s</sub> is the active component, as the tiny Rpf<sub>smeg</sub> consists of little else but this domain. Studies of the activity of Rpf<sub>smeg</sub> are described in Chapter 5.

In this chapter it was also demonstrated that soluble RpfB and RpfE can be obtained by using the Origami B (DE3) pLysS host, lower induction temperatures and lower concentrations of inducer - albeit in amounts approximately ten times less than was achieved using urea as a denaturant. Applying these modifications to a large-scale culture, extraction and purification, the yield of purified RpfE was approximately 2.24 mg from 2 l – approximately 1 mg per litre of culture. These concentrations were unfortunately still barely sufficient for successful NMR spectroscopy. In the meantime, structural biologists working at Birkbeck college were achieving better protein yields and solubility using a new construct consisting of the core Rpf domain alone, and were optimistic of obtaining high enough concentrations for crystallisation. Therefore this aspect of the work was not taken further. See Cohen-Gonsaud *et al* (2005) for the eventual results of this structural work.

In order to further improve the solubility and yield of RpfE, a possible next step would have been to clone the gene into the pET-43.1 vector to create a fusion protein with

NusA, as was done with Rpf<sub>smeg</sub>. This, combined with a slow induction at low temperature, might have resulted in better solubility.

Although large quantities of protein were not obtained, nonetheless an improvement in solubility had been achieved, particularly in the case of Rpf<sub>smeg</sub>. The next step was to test the biological activity of the Rpfs, so that they could be shown to be suitable for use in further experiments. In order to do this, suitable bacterial cultures containing cells in a dormant state were required. Chapter 5 will describe how these cell cultures were set up and investigated; and the various activity testing methods which were tried and developed.

## **CHAPTER V**

### **RESULTS: ACTIVITY TESTING OF Rpf USING CULTURES OF *MYCOBACTERIUM BOVIS* BCG AND *MICROCOCCUS LUTEUS***



## 5.1 Introduction

Activity testing of Rpf has proved to be a problematic area. The activity testing methods in use when this project was started were laborious to set up, lengthy (5-29+ days depending on the organism used) and not always reliable. Testing was, however, always necessary, since even in experienced hands, batches of Rpf are found to vary in their activity, from highly active to no activity at all, and furthermore loss of activity occurs rapidly, within a few days or less, even when the proteins are stored at  $-20^{\circ}\text{C}$  in 50% glycerol. All the activity tests employed required cells from dormant cultures of responsive organisms to demonstrate Rpf activity. As the activity of *M. tuberculosis* Rpf was to be studied, the ideal test organism would have been *M. tuberculosis* itself. However, at the laboratory in the University of Wales, Aberystwyth, where this part of the project was done, only Containment Level 2 facilities were available, so *M. bovis* BCG was used as a test organism instead. This has been shown to be an excellent model for dormancy in *M. tuberculosis*. (Lim 1999, Shleevea 2002).

In order to be able to perform activity assays using *M. bovis* BCG, it was first necessary to establish and characterise dormant cultures of this organism. An important objective was to establish cultures under standard conditions that would produce reproducible results as an experimental system. This would permit the observation of the growth characteristics of the model, and the response to Rpf, over a six month period. Section 5.4 of this chapter describes these experiments; section 5.5 describes the ensuing use of these cultures for testing activity of Rpf.

Subsequently, *Micrococcus luteus* was used for activity testing, because of its more rapid growth. In sections 5.6 and 5.7 activity testing assays using *M. luteus* cultures will

be described; in particular, assays developed using measurement of ATP bioluminescence of cultures as an indicator of Rpf activity.

## **5.2 *In vitro* models of dormancy in *M. tuberculosis***

The best known *in vitro* model of dormancy in *M. tuberculosis* is the Wayne model, which relies on oxygen depletion to induce the dormant state. In the original Wayne model (Wayne 1976) cultures of *M. tuberculosis* were sealed in tubes and left to stand without agitation. The cultures became gradually more anaerobic and those organisms which sank to the bottom of the tubes first experienced more anaerobic conditions than those near the surface where there was a headspace of air. Wayne realised that this resulted in a heterogenous bacterial population and modified his model to involve agitation of the sealed vessels with a magnetic stirrer, so that all the organisms experienced anaerobiosis at the same time (Wayne 1996). Various headspace: broth ratios from 1:1 to 1:4 were used: the less the headspace the less the optical density the culture reached.

Wayne described dormancy as a state of 'non-replicating persistence' (Wayne 1996) and distinguished between two persistent states which ensued with his model: non-replicating persistence 1 (NRP1) and NRP2. NRP1 is a microaerophilic state, in which the organisms increase in size but do not replicate. When oxygen saturation reaches <0.06% NRP2 follows, which is an anaerobic state in which growth ceases altogether and alterations in drug susceptibility are observed. For example, in anaerobic conditions the bacilli lose their susceptibility to ciprofloxacin and isoniazid and have reduced susceptibility to rifampicin. They become susceptible to metronidazole, which has no activity on aerobic cultures of *M. tuberculosis*. Wayne's published experiments did not continue much beyond 16 days, by which time cultures were declared 'dormant' as evidenced by loss of turbidity and cfu/ml.

Reintroduction of oxygen led to resumption of growth, which proceeded in an apparently regulated, synchronous fashion, with RNA synthesis preceding DNA synthesis. This model has been claimed to have clinical correlation with human lung cavities in which anaerobic conditions might prevail, and in which acid-fast bacilli may be microscopically visible but not culturable.

Dick's group in Singapore used the Wayne model with *M. bovis* BCG (Lim 1999). The organisms became 'dormant' after 20 days (ie no further increase in turbidity or cfu/ml) and they concluded that the 'dormancy response of BCG is almost identical to that of *M. tuberculosis*'. The same model was also used to study *M. smegmatis* (Dick 1998).

Other approaches have used models based on low pH (Loebel 1933, Heifets 1992), complete anaerobiasis (Loebel 1933, Cunningham 1998), and nutrient starvation (Loebel 1933, Smeulders 1999, Betts 2002).

For the purposes of this work, as discussed in section 1.2, dormancy was defined as being in 'a state of low metabolic activity and unable to divide or form a colony on an agar plate without a preceding resuscitation phase'. In other words, dormancy is a state in which cells are present which cannot grow upon subculture on solid or liquid medium unless resuscitated, for example in this case by incubation in a suitable liquid medium containing Rpf. Although the models investigated were, like Wayne's, based on oxygen deprivation, a much greater headspace was used in the vessels (headspace: broth ratio 5:1), so that oxygen depletion would be more gradual, giving the organisms sufficient oxygen to grow to a higher density, at which nutrient depletion would normally also occur. It was postulated that gradual adjustment to the lowering oxygen tensions at a higher cell density might result in a more reproducible induction of dormancy. The cultures were studied over the course of 6 months, a much longer time span than that

described by Wayne, and in a minimal medium, rather than Wayne's Dubos Tween-albumin. The presence of cells requiring Rpf to resuscitate them before they could divide was demonstrated by estimating the Most Probable Number (MPN) of viable cells in a liquid medium in the presence and absence of Rpf.

### **5.3 Most Probable Number tests**

Most probable number testing is a method of obtaining viable counts depending on dilution to extinction. A sequence of dilutions of a culture is made. At the levels where the concentration of cells ranges from 10-0.1/ml there is a statistical probability that progressively fewer of the subcultures will grow. From the distribution of tubes showing growth at three such levels of dilution, the most probable number of bacteria in the preceding dilution can be obtained from tables (see Appendix v). Though the method is convenient it is relatively inaccurate as it is based on statistical probabilities: however where the viable count needs to be known within a factor of about two it is adequate.

### **5.4 Establishing a dormancy model in *M. bovis* BCG**

#### *5.4.1 Method: setting up of cultures*

Two types of culture vessel were used – airtight glass test tubes with a rubber seal, and 500 ml glass flasks, also with an airtight rubber seal. The headspace:broth ratio was 5:1 in both flasks and tubes. The test tubes were kept in a 37°C incubator and remained stationary, whilst the flasks were incubated and constantly agitated in a waterbath, also at 37°C. Both contained Sauton's medium inoculated with dilutions of a 6 day old culture of *M. bovis* BCG. The flasks contained 200 ml cultures, and the tubes 2 ml cultures. The dilutions were calculated to give a starting cell count of 2 cfu/ml in the

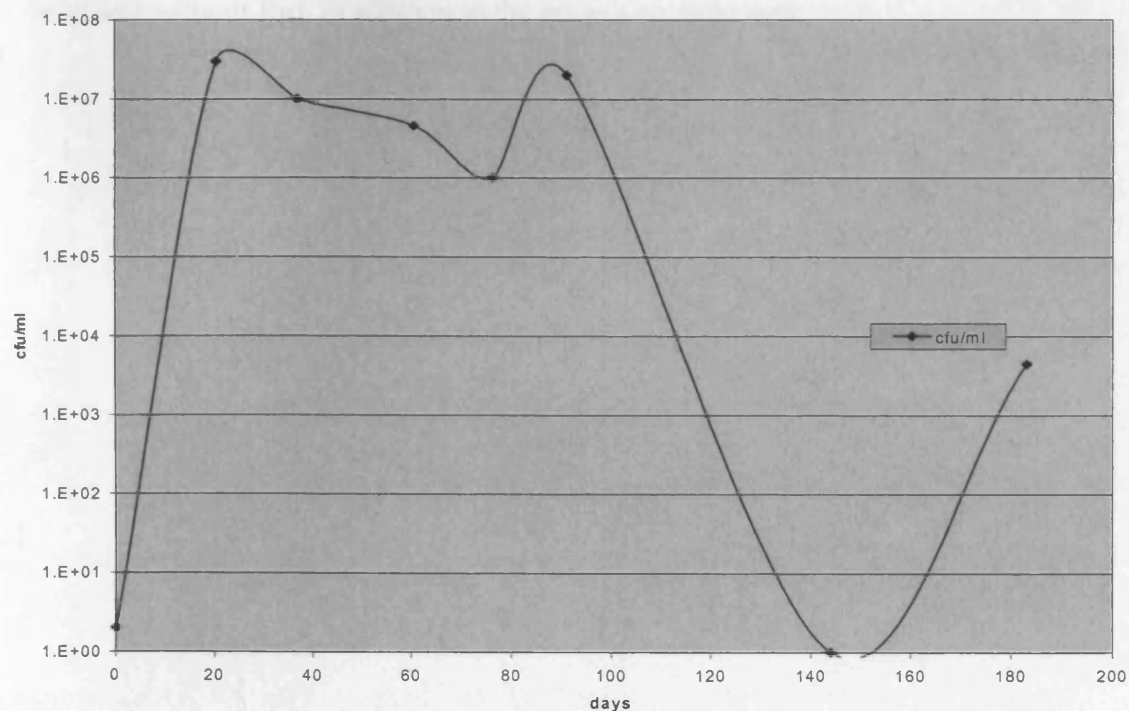
tubes, and 20 cfu/ml in the flasks. For a full description of the methodology used see Methods section 2.4.

#### 5.4.2 Results

##### 5.4.2.1 Tubes – viable counts

The tubes were sampled at intervals and plated out onto Middlebrook agar to determine the viable count. The results are shown in Figure 5.1. These demonstrate that in the first 20 days of the tube model there was rapid growth of the cultures. Thereafter the bacterial population began to decline, reaching a nadir of  $10^6$  cfu/ml on solid agar at 76 days (two and a half months). There then appeared to be slight recovery but by 144 days (about 5 months) there was no growth at all on solid agar.

**Figure 5.1** Viable count of *M. bovis* BCG cultured in Sauton's medium in stationary tubes, subbed onto solid agar plates over time.

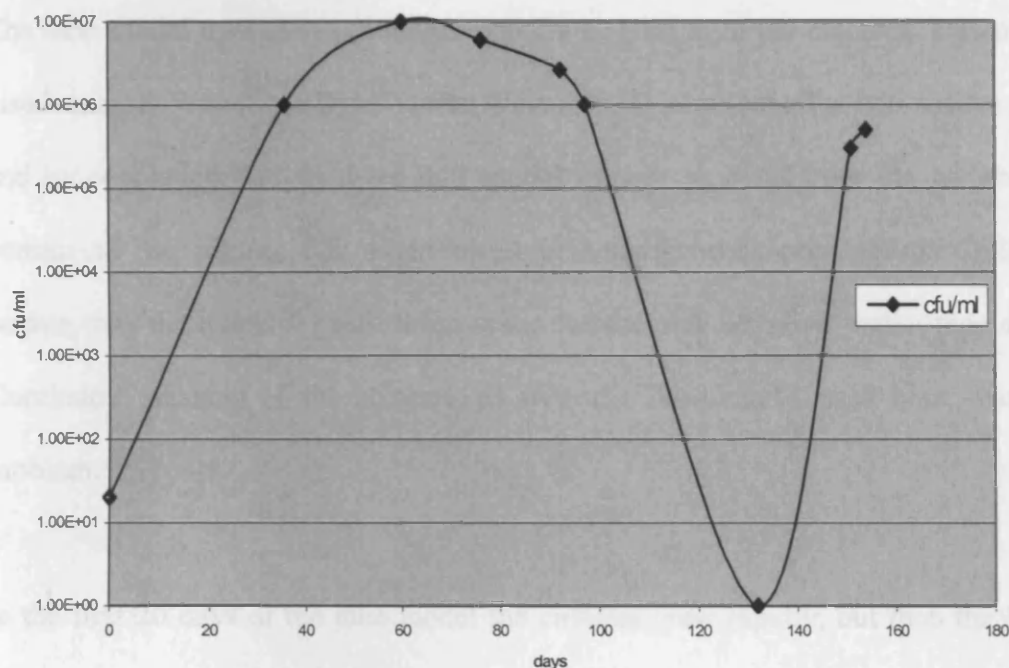


#### 5.4.2.2 *Flasks – viable counts*

The flasks were sampled at intervals and plated out onto Middlebrook agar to determine the viable count. Most Probable Numbers in liquid medium (Sauton's) with and without Rpf was also measured later in the course of the experiments. The results are shown in Figure 5.2. These show that bacteria in the flask model grew initially up to  $\sim 10^7$  cfu/ml on solid agar by 60 days (2 months). There followed a decline to  $10^6$  cfu/ml by 3 months, after which there was a rapid drop to zero viability on solid agar at 4.5 months. However the organisms then started to recover again, for reasons which are uncertain (see Discussion section 5.4.4). Viable counts in liquid media (by MPN) were consistently higher by at least an order of magnitude, whenever they were measured. Addition of Rpf to the liquid cultures at 111 days resulted in a four-fold increase in viable counts as measured by MPN, but this may possibly be attributable to experimental error since the reverse was found at 151 days. At 168 days (5½ months) addition of Rpf to the liquid cultures resulted in a more convincing eight-fold increase in the viable counts. Appendix vi shows the results of the MPN tests in liquid medium with and without Rpf, in addition to the growth on solid agar.

**Figure 5.2**

Viable count of *M. bovis* BCG cultured in Sauton's medium in constantly agitated flasks, subbed onto solid agar plates over time.



#### 5.4.4 Discussion – dormancy model in *M. bovis* BCG

The model of dormancy used here is based on gradual oxygen depletion. The large headspace means that initially cells are well aerated. Later on, microaerophilic conditions ensue and only when all the oxygen in the headspace has been used do the organisms become anaerobic. The main difference between this and the Wayne model is that the headspace in the Wayne model is much smaller, resulting in more rapid oxygen depletion: other differences are the use of a minimal medium (Sauton's) instead of Dubos Tween-albumin broth in the Wayne model, and a slightly lower pH in the Wayne model (~6.6, as opposed to 7.0 in this model). The Wayne model may be open to criticism because when inoculated into fresh medium the cells began to grow immediately, and apparently normally. This raises the possibility that they were not 'dormant', but simply that some of the population had died resulting in the observed

decrease in cfu/ml. This criticism cannot be levelled at the model used here, in which dormancy had to be demonstrated by the presence of organisms which could only grow after resuscitation by Rpf.

The tube model used did not include shaking or stirring of the cultures. This may be a disadvantage. Wayne's original work (Wayne 1976) also looked at non-shaken cultures and he concluded that, as there is a spatial oxygen gradient from the surface to the bottom of the culture, this might result in heterogeneous populations. Cells at the bottom may die quickly, whilst those at the surface may be 'alive' rather than dormant. Continuous shaking of the cultures, as with the flask model used here, avoids this problem.

In the first 20 days of the tube model the cultures grew rapidly, but then the bacterial population began to decline. A slight recovery in the counts was observed at about 3 months but by 5 months there was no growth at all on solid agar.

Bacteria in the flask model also grew initially, up to  $\sim 10^7$  cfu/ml on solid agar by 60 days (2 months). The peak of the curve was reached about a month later than with the tube cultures. There followed a decline to  $10^6$  cfu/ml by 3 months, after which there was a rapid drop to zero viability at 4.5 months. However the organisms then started to recover again, a finding replicated in experiments by others in work published subsequently (Shleeva *et al* 2002). A similar phenomenon was also noted in cultures of certain mutants of *M. smegmatis* (Keer, Smeulders & Williams 2001), which recovered culturability after 8-15 days of oxygen starvation. This group's interpretation was that either the population had begun to die but then regrown using an unknown fermentation



pathway, or, that the population had become transiently dormant or unculturable and then undergone resuscitation as conditions changed in stationary phase.

Most Probable Number determinations (in liquid broth), performed on the culture from the flask at 111, 132 and 151 days, consistently showed greater viability than the viable count on solid agar. Furthermore it was possible to demonstrate an increase in cfu/ml by the addition of Rpf (these results are described in more detail in section 5.5). This means that possibly by 4 months, and more certainly by 5½ months, 'dormant' cells, defined as those rescusitatable by Rpf, were present.

Further experiments would be required to determine whether dormant cells, as evidenced by action of Rpf on MPN counts, are present earlier than 4-5½ months. At 3 months, a drop in cfu was already apparent. It would be helpful to be able to produce dormant cells in the minimum time possible. Experiments to compare the Wayne model with this one would also be helpful. Perhaps the smaller headspace of the Wayne model has the effect of causing oxygen depletion to occur too quickly: cells may die because they cannot adapt to oxygen depletion under conditions when nutrients are in excess. Alternatively, the Wayne model might simply have the effect of bringing about a state of dormancy more rapidly.

## **5.5 Activity testing of mycobacterial Rpf by MPN and concentration-activity tests on *M. bovis* BCG**

Two types of experiment were used to test the effect of Rpf on dormant cultures of *M. bovis* BCG. They took the form of conventional MPN tests, or, a 'concentration-activity test'. These two tests are described fully in Methods 2.4.2, but are briefly

reviewed below. Both used cultures of *M. bovis* BCG, produced as described in 5.4, which had been cultured in Sauton's medium for between 3 and 6 months.

### 5.5.1 *Methods*

#### 5.5.1.1 *MPN test*

Rpf at a picomolar concentration was used. Dormant *M. bovis* BCG cultures were used at a range of dilutions from  $10^{-2}$  –  $10^{-7}$ . A grid of tubes was set up to test the Rpf against each dilution of *M. bovis* BCG, using five replicates of each tube. A tube became 'positive' when growth was visible at the bottom of the tube using the naked eye (growth appeared as a white 'button' at the bottom of the tube). The highest dilution of *M. bovis* BCG culture which was able to grow with and without addition of Rpf was recorded.

#### 5.5.1.2 '*Concentration-activity*' test

This test differed from the MPN test in that it was the concentration of Rpf which was varied, in tubes with a constant dilution of dormant *M. bovis* BCG culture (rather than varying the dilution of *M. bovis* BCG with a constant concentration of Rpf). This allowed testing for the most active concentration of a particular batch of Rpf. Testing was set up using a  $10^{-3}$  dilution of dormant *M. bovis* BCG cultures, with Rpf added at a range of concentrations between nanomolar and femtomolar. Five replicate tubes were set up for each concentration of Rpf.

### 5.5.2 *Results*

#### 5.5.2.1 *Results RpfA: concentration-activity test*

The activity of a sample of RpfA prepared 24 h earlier and stored in glycerol (50:50 v/v) at -20°C was tested. The stock solution of RpfA was diluted and tested for activity at concentrations between 100 pM and 1 aM. The different concentrations of Rpf were

added tubes containing 1:1000 dilutions of a 5-month old dormant culture of BCG incubated for 5 months. Tubes were read daily to look for growth, appearing as a 'button' at the bottom of the tube, detectable by the naked eye. It was necessary to check tubes daily since once all the tubes displayed growth the result would be meaningless. The test relied on the rapidity of growth in each tube relative to the others. For clarity, results are shown in tabulated form (table 5.1). A '+' sign represents visible growth in that tube.

**Table 5.1**

Result of RpfA concentration-activity test at 9 days:

Concentration of Rpf	1	2	3	4	5
100 pM	+	+	+	+	+
1 fM	-	-	-	-	-
10 fM	-	-	-	-	-
100 fM	+	-	-	-	-
1 aM	-	-	-	-	-

At 14 days all the tubes displayed visible growth.

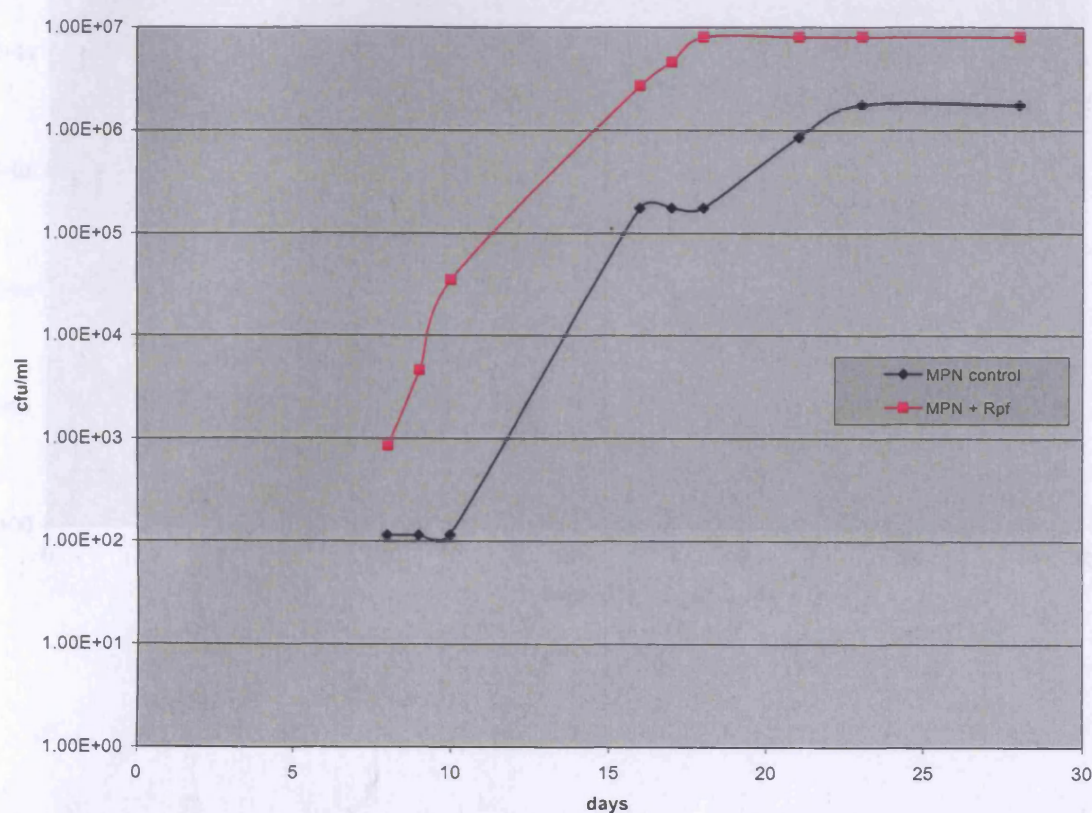
This result shows that in tubes containing RpfA at a concentration of 100 pM, growth was seen when there was no growth in tubes containing lower concentrations of RpfA. Growth occurred with RpfA at 100pM 5 days before it occurred in most of the other tubes, indicating that RpfA has activity at this concentration. (No control without RpfA was set up on this occasion. This should have been done to exclude the presence of any inhibitory substances in the RpfA stock. However such inhibitors would have been

present in ever decreasing concentrations in each dilution. Since the tubes with the most concentrated RpfA grew fastest, an inhibitory effect in the other tubes is very unlikely).

#### 5.5.2.2 Results RpfB: MPN test

To test the activity of RpfB the MPN method was used rather than a concentration activity test. An RpfB concentration of ~25 pM was used. Dormant BCG cultures (5.5 months old) were used at dilutions of  $10^{-2}$ – $10^{-7}$ . The protein was used 3 days after production. The data are summarised in the growth curves seen in Figure 5.3. The final MPN without Rpf was  $1.75 \times 10^6$ . The MPN in the presence of Rpf was  $8 \times 10^6$ , i.e. increased by about a factor of four. The results show both a reduced lag phase and a higher MPN (growth at greater dilutions) with 25 pM RpfB than without.

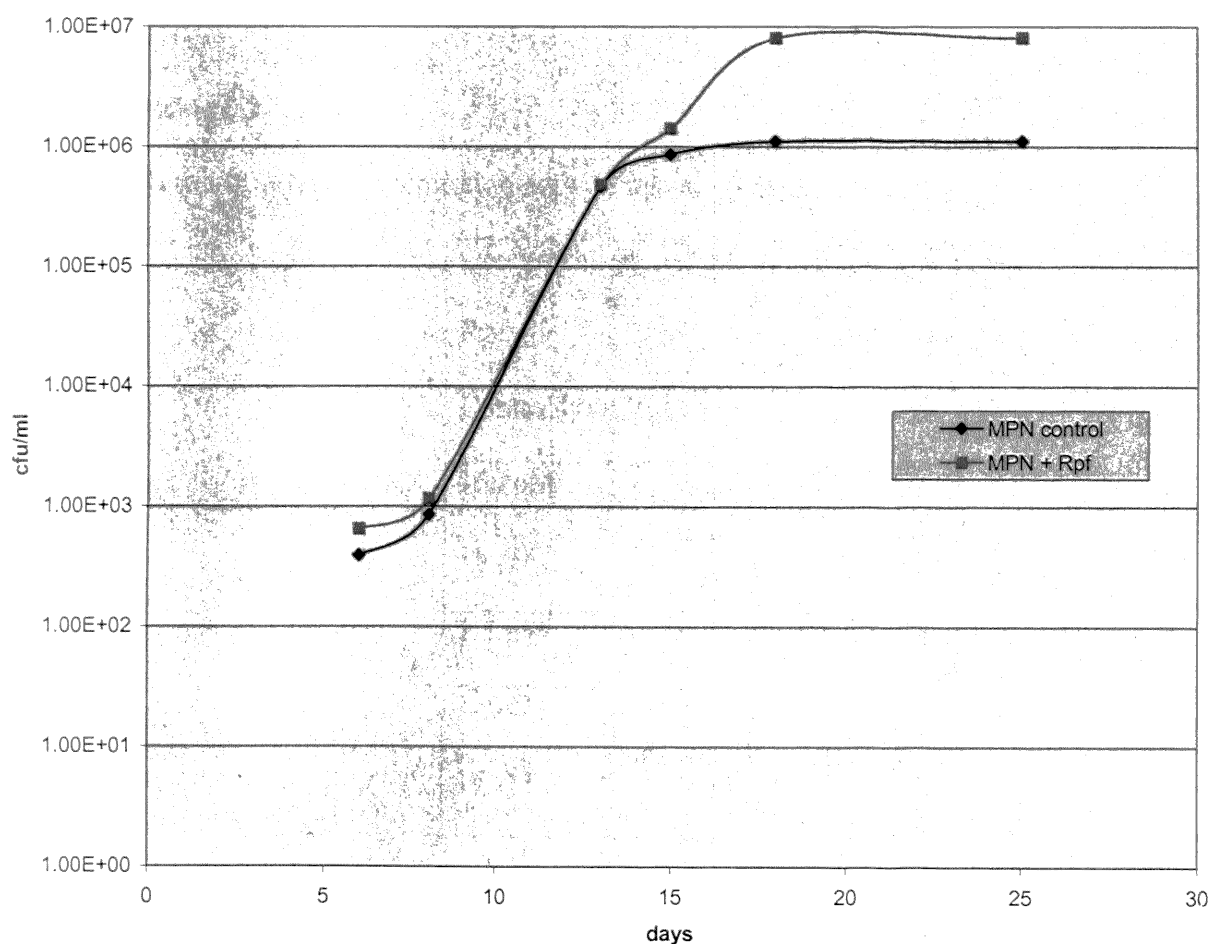
**Figure 5.3** Growth curve of *M. bovis* BCG (5.5 month-old culture) with and without RpfB. The effect of adding 25pM RpfB to the culture is shown in pink. The negative control (without Rpf) is in blue.



### 5.5.2.3 Results *RpfC*: MPN test

An MPN assay was set up using 10 pM concentration of fresh *RpfC*. The data are summarised in the growth curves in Figure 5.4. The MPN without *Rpf* was  $1.1 \times 10^6$  cfu/ml; the MPN with *Rpf* was  $8 \times 10^6$  cfu/ml, a difference of a factor of  $\sim 8$ . *RpfC* has biological activity at the 10 pM concentration used. In contrast to the result with *RpfB* however a reduced lag phase was not evident, indicating that these two *Rpfs* have different effects.

**Figure 5.4** Growth curve of *M. bovis* BCG (5.5 month old culture) in Sauton's medium with and without *RpfC*. The effect of adding 10pM *RpfC* to the culture is shown in pink. The negative control (without *Rpf*) is in blue.



#### 5.5.2.4 Results RpfD: concentration-activity test

RpfD in concentrations ranging from 15 fM to 150 pM (see Table 5.2) were used, with a  $10^{-4}$  dilution of cells from a dormant BCG culture. Negative control tubes contained no Rpf. The protein was 24 h old when used. At 10 days, all five tubes containing RpfD at a concentration of 1.5 pM were displaying growth, compared with two or three tubes only for the other concentrations (Table 5.2). By 13 days all the tubes showed visible growth. This result may indicated some activity of RpfD at a concentration of 1.5 pM, compared to the negative control tubes, but the possibility that the effect is due to random variation cannot be dismissed. Further experiments, using a larger number of tubes with containing 1.5pM RpfD, and more negative controls, would be required to conclusively demonstrate a significant difference.

**Table 5.2**

Result of RpfD concentration-activity test at 10 days:

Concentration of Rpf	1	2	3	4	5
Controls (no Rpf)	+	+	-	-	+
150 pM	+	-	-	+	-
15 pM	+	-	-	-	+
1.5 pM	+	+	+	+	+
150 fM	-	+	+	+	-
15 fM	+	+	+	-	-

#### 5.5.2.5 Results RpfE: concentration-activity test

The activity test was set up using RpfE in pM concentrations.  $10^{-4}$  concentrations of BCG from a 5 month old dormant culture were used. The RpfE was 24 h old. Tubes

were read daily, however the experiment did not appear to show any activity of RpfE, at any concentration. As RpfE is expressed at high concentrations, with a heavy band visible on SDS-PAGE, aggregation may have been occurring, with consequent loss of activity. In Chapter 4 RpfE was induced using lower concentration of IPTG, at lower temperatures and longer induction periods, in an attempt to improve its solubility.

#### *5.5.2.6 Results Rpf<sub>smeg</sub>: concentration-activity tests and MPN tests*

The activity profile of Rpf<sub>smeg</sub> had not been tested before: its expression had been too poor to produce sufficient quantities for testing. However, once a soluble form had been produced, as described in Chapter 4, using the NusA fusion protein and Origami (DE3) host, much greater quantities were obtained, so that the activity could be tested.

Four batches of Rpf<sub>smeg</sub> were tested for activity and the results are described below. The results of the Rpf<sub>smeg</sub> activity testing are summarised in Table 5.6.

##### *5.5.2.6.1 Experiment 1: concentration-activity test*

Rpf<sub>smeg</sub> was used two days after it had been purified and stored at -20°C. A 10<sup>-3</sup> dilution of a dormant BCG culture (age 3 months) was used to find the most active concentration of Rpf<sub>smeg</sub>. Rpf<sub>smeg</sub> activity was tested at concentrations between picomolar and femtomolar. Table 5.3 shows the results at 15 days.

**Table 5.3**

Growth of dormant cultures of *M. bovis* BCG in the presence of Rpf<sub>smeg</sub> in concentrations between pM and fM, after 15 days. '+' indicates visible growth of BCG in the tube:

Rpf dilution	1	2	3	4	5
No Rpf -control	-	-	-	-	+
18 pM	-	-	+	-	+
1.8 pM	-	-	-	-	-
180 fM	-	-	-	-	-
18 fM	+	+	+	+	-
1.8 fM	+	+	+	+	+

After 19 days growth was visible in all the tubes. These results suggested that, compared with tubes containing no Rpf, BCG in tubes with 18 fM and 1.8 fM of Rpf<sub>smeg</sub> grew more quickly, i.e. the protein had biological activity at 1.8 fM.

The experiment was repeated, using the same Rpf after storage for 2 weeks at -20°C. Rpf<sub>smeg</sub> no longer appeared to have significant activity.

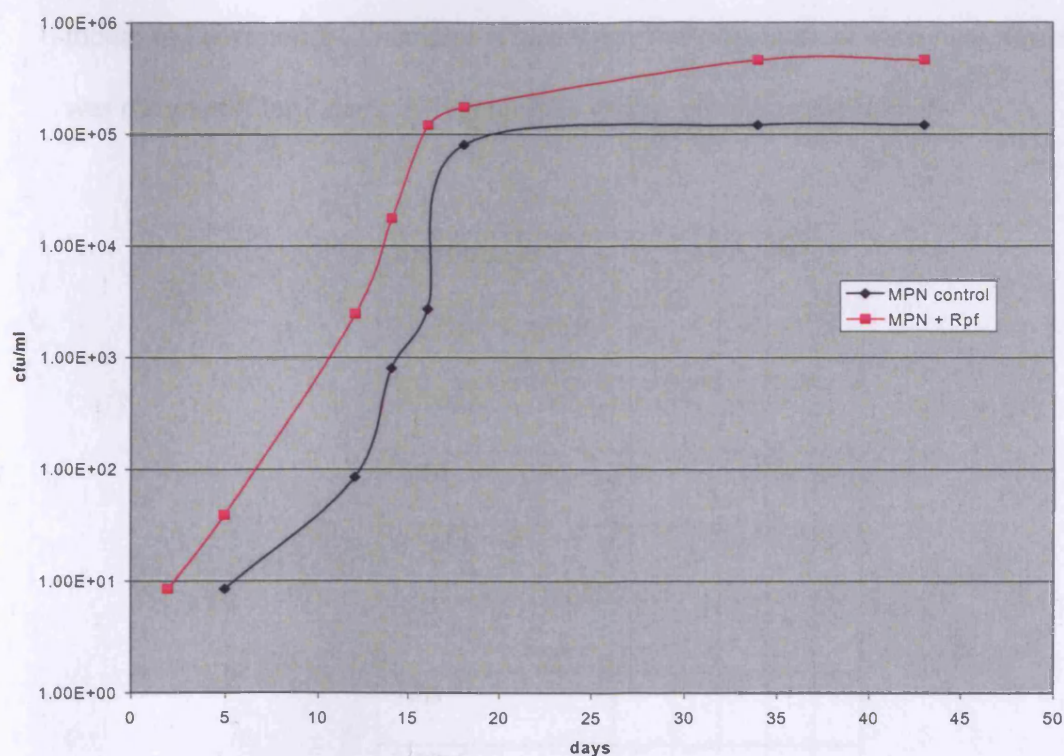
#### 5.5.2.6.2 Experiment 2: MPN test

For MPN testing dormant BCG cells (culture age 4 months) were used at dilutions of 10<sup>-1</sup>-10<sup>-6</sup>. One-day old Rpf<sub>smeg</sub> was used at a concentration of ~50 fM. The tubes were read daily until there was no further growth in any tube for at least 10 days. New growth in tubes ceased after 34 days. Throughout the experiment, the number of positive tubes with Rpf<sub>smeg</sub> exceeded the number without Rpf<sub>smeg</sub>, usually by a difference of one row; ie there was visible growth in tubes containing ten-fold less inoculum when Rpf<sub>smeg</sub> was



present. Towards the end of the experiment, the tubes without Rpf started to catch up as growth had stopped in those with Rpf: but even so the final result showed a greater MPN when estimated in the presence of Rpf. Using MPN tables (Appendix 3) it was estimated that without Rpf, the undiluted culture of dormant BCG appeared to contain  $1.2 \times 10^5$  viable organisms. With Rpf, the number of viable organisms detected was approximately four times this amount, at  $4.6 \times 10^5$ . Growth curves for these BCG cultures with and without Rpf<sub>smeg</sub> are shown in Figure 5.5.

**Figure 5.5** Growth curve of *M. bovis* BCG (4 month old culture) in Sauton's medium with and without *M. smegmatis* Rpf. The effect of adding 50 fM *M. smegmatis* Rpf is shown in pink. The negative control (without Rpf) is in blue.



#### 5.5.2.6.3 Experiment 3: MPN and concentration-activity tests

A third batch of Rpf<sub>smeg</sub> was produced, and was 3 days old when it was used for this experiment. MPN determination was carried out using this protein at a concentration of

about 70 fM, and a 5-month old dormant BCG culture. A concentration activity test was also set up. Neither of these tests revealed any difference between the cultures containing Rpf<sub>smeg</sub> and the controls.

#### 5.5.2.6.4 Experiment 4: MPN and concentration-activity tests

A fourth batch of Rpf<sub>smeg</sub> was produced. Both concentration activity testing and MPN determination was carried out, when the protein was one day old. The concentration activity test used a 10<sup>-3</sup> dilution of a 6-month old dormant BCG culture, and concentrations of Rpf<sub>smeg</sub> in the pM – fM range. Results are shown in table 5.4.

**Table 5.4**

Concentration-activity testing of Rpf<sub>smeg</sub> in pM - fM concentrations, in tubes of a 6-month old dormant BCG culture. There were five replicates of each tube. Result shown was obtained after 7 days. A '+' signifies visible growth in the tube.

Rpf dilution	1	2	3	4	5
No Rpf- control	-	-	-	-	-
70pM	+	-	-	-	-
7pM	+	+	+	-	-
0.7pM	-	-	+	-	-
70fM	-	+	-	-	-
7fM	-	-	-	-	-
0.7fM	-	-	-	-	-
70aM	-	-	-	-	-

After 9 days most tubes contained visible growth. This result may be indicative that the Rpf was active at 7pM but without further replicates it would not be possible to rule out that it is a random effect. A follow up experiment would be required using more replicates at 7pM rather than systematic ones as performed.

**Table 5.5**

Summary of results of Rpf<sub>smeg</sub> activity testing:

Batch of Rpf <sub>smeg</sub>	Stock concentration	Age of stock	Type of test	Rpf <sub>meg</sub> final concentration (MPN)	Increase in viable count with Rpf (MPN)	Most active concentration (concentration -activity test)
1	1.8uM	48h	Concentration -activity	-	-	1.8fM
1	1.8uM	2 weeks	Concentration -activity	-	-	No activity
2	500nM	24h	MPN	50fM	4-fold increase	-
2	500nM	24h	MPN	50aM	No activity	-
3	7uM	72h	MPN	70fM	No activity	-
4	7uM	24h	Concentration -activity	-	-	7pM

### 5.5.3 Summary of results of activity testing of mycobacterial Rpfs on dormant cultures of *M. bovis* BCG

Four of the five Rpf-like proteins of *M. tuberculosis* (recombinant versions) showed biological activity in these tests: this was least marked for RpfD (Rv2389c). RpfE (Rv2450c) did not show any biological activity on this occasion, possibly because of aggregation.

Of four batches of Rpf<sub>smeg</sub> produced without the use of urea, three displayed biological activity, using either MPN or concentration-activity testing (see Table 5.5). No activity was seen after the protein had been stored for 2 weeks. Concentration-activity testing indicates a range of activity at between pM and fM concentrations. This protein consists only of an Rpf-like domain (shared by all known members of the Rpf protein family) together with a secretory signal sequence. These experiments therefore confirm that the biological activity of these proteins resides in their Rpf-like domain. It was hoped that the Rpf<sub>smeg</sub>, which unlike the other Rpfs was produced without the use of urea, might be more reliably active and stable. This may be the case, but was not conclusively demonstrated: more experiments would be needed to prove this.

## **5.6 Activity testing by measuring OD<sub>600</sub> of *M. luteus* cultures in microtitre plates**

The length of time to a result, and the laboriousness of the methods, limited the usefulness of tests on cultures of *M. bovis* BCG for testing activity of Rpfs. A more rapid and reliable activity test was considered a priority for this project, so that later experiments could rely on known active protein being available. Since it is known that *M. tuberculosis* Rpfs are active on *M. luteus* cells, which are relatively rapidly growing, tests based on this organism were investigated.

Attempts were made to measure the activity of Rpf on *M. luteus* by measuring change in OD<sub>600</sub> in cultures in microtitre plates (as described by Mukamolova 2002a and 2002b). These experiments were carried out using both *M. luteus* Rpf and Rpf<sub>smeg</sub>. See Methods 2.5.1 for a full description of the assay. The test employed a low inoculum (10<sup>3</sup> cfu/ml) of stationary phase culture of *M. luteus* in lactate minimal medium (LMM),

as these are the experimental conditions in which the best activity of Rpf has been demonstrated. The inoculum was passed through a needle ten times but despite this clumping of cells remained and this tended to result in uneven growth of *M. luteus* in random wells. To overcome the problem of clumping, the *M. luteus* cells were filtered through a 5 µm filter to try to remove clumps, and to balance the effect of filtering resulting in loss of cells, a tenfold higher concentration of *M. luteus* cells was used: however after filtering no growth was obtained in any of the wells. On other occasions, even with unfiltered cells, no growth occurred in any well at 10 days, either with or without Rpf. To check that the LMM could support bacterial growth properly, the *M. luteus* suspension was simultaneously inoculated into a flask containing the LMM, to give the same final cell density: this grew well overnight. Quantitative counts of the inoculum of *M. luteus* were also made on LB plates: this showed that ~500 cfu had been inoculated into each well, considered a sufficient amount. Furthermore the cultures in the microtitre plates required so much manipulation during the complex setting-up process that there was a high risk of contamination, even though the plates were set up in a Class 2 cabinet to prevent this.

None of the experiments using this assay were successful - the results are summarised in Table 5.6.

**Table 5.6**Summary of results of activity testing using *M. luteus* in microtitre wells

Experiment number	Rpf used	Age of Rpf	Filtered/ unfiltered	Starting inoculum (cfu/ml)	Result
1	Rpf	fresh	unfiltered	$10^3$	Clumping in random wells
2	Rpf	1 week	filtered	$10^4$	Contaminated
3	Rpf <sub>smeg</sub>	fresh	filtered	$10^4$	Contaminated
4	Rpf	1 week	filtered	$10^4$	No growth any well at 2 weeks
5	Rpf	1 week	filtered	$10^4$	No growth any well at 2 weeks
6	Rpf <sub>smeg</sub>	1 week	unfiltered	$10^3$	No growth any well at 10d
7	Rpf	1 week	unfiltered	$10^3$	No growth any well at 10d

These results were disappointing, though not entirely unexpected, since experienced workers reported difficulty in obtaining reproducible data from one experiment to the next, with this model (see Discussion). The tendency of the cells to clump in culture, and the very low inocula required to demonstrate the effect of Rpf, both contributed to the problems of poor reproducibility. No further experiments were done using this method.

### 5.7 Activity testing using ATP bioluminescence of *M. luteus* cultures

ATP-bioluminescence is a well-established technique which has been used as a measure of bacterial counts, particularly in food and environmental microbiology. Luciferase from the American firefly catalyses the following reaction:



The amount of ATP present can be calculated from the amount of light emitted and is proportional to the bacterial count/ number of cells present. Even extremely low concentrations of ATP can be measured. Thus by measuring the ATP content of a culture, the problem of cells clumping could be overcome. An assay was therefore developed which could measure the difference in concentration of ATP between cultures of *M. luteus* grown with and without Rpf in LMM. The ATP content was measured by bioluminescence in a luminometer. Two assays were developed: an overnight test, which had greater discrimination between batches, and a rapid four hour test, which could more quickly demonstrate the presence or absence of active protein, but which was less discriminating between batches.

### 5.7.1 Overnight ATP bioluminescence activity test

#### 5.7.1.1 Method: overnight ATP bioluminescence activity test

The method is described in detail in section 2.5.3.1 but to summarise, stationary phase cultures of *M. luteus* were diluted a thousand-fold in LMM and incubated with Rpf for 16-18 hours. The inoculum was equivalent to  $10^6$  cfu/ml, higher than that in 5.6. Growth was slight under these conditions with turbidity normally not visible to the naked eye, though sometimes small clumps of cells were seen.

All the experiments were done using RpfB. RpfB was expressed from both clone CA1002 (pET19b vector and HMS174 host) and clone AD5 (pET19b vector and Origami B (DE3) pLysS host) to see which expression system produced the most active recombinant protein. RpfB was chosen because in contrast to the other *M. tuberculosis* Rpfs, its activity profile had not previously been measured against *M. luteus*

(Mukamolova 2002b) - although it was known to have activity in cultures of *M. bovis* BCG - and because it is relatively easy to prepare reasonable concentrations of the recombinant protein. The positive control was *M. luteus* culture supernatant, obtained as described in 2.5.3.1; the negative control was a dilution of elution buffer in LMM, equivalent to the dilution of Rpf used.

After spinning down the cultures and extracting ATP, the ATP bioluminescence of the cultures and of ATP standards was measured in a luminometer. A standard curve was drawn using the results for the ATP standards, on a log-log graph, to correlate bioluminescence with ATP concentration. The ATP concentrations in the cultures could then be read off the graph.

#### *5.7.1.2 Results: overnight ATP bioluminescence activity test*

These results are summarised in Table 5.7 and are described below.

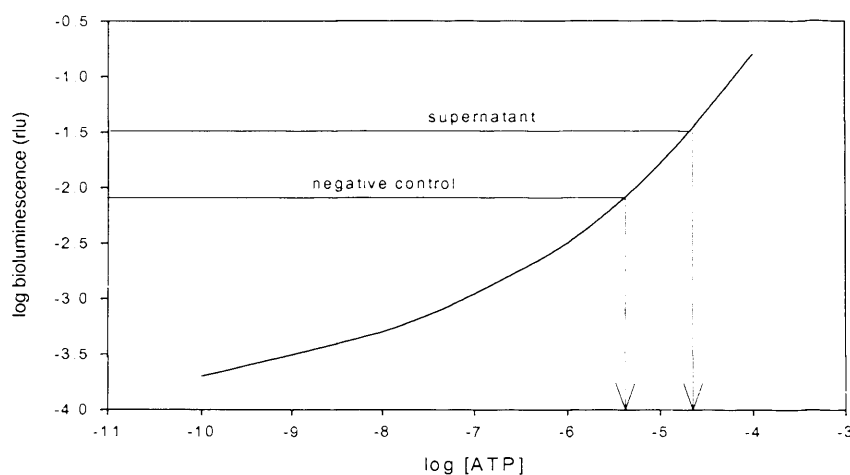
##### *5.7.1.2.1 Results: positive control*

A range of activity of the positive control was found, with from 2.4 – >35 x greater concentrations of ATP than the negative control. The >35 x greater concentration of ATP was found after culture supernatant was used immediately; the other results (2.4 x, 4.5 x and 5 x greater concentration of ATP) were obtained after the supernatant had been kept on ice for 30-60 min. Examples of these results are shown as graphs in figures 5.6 and 5.7.



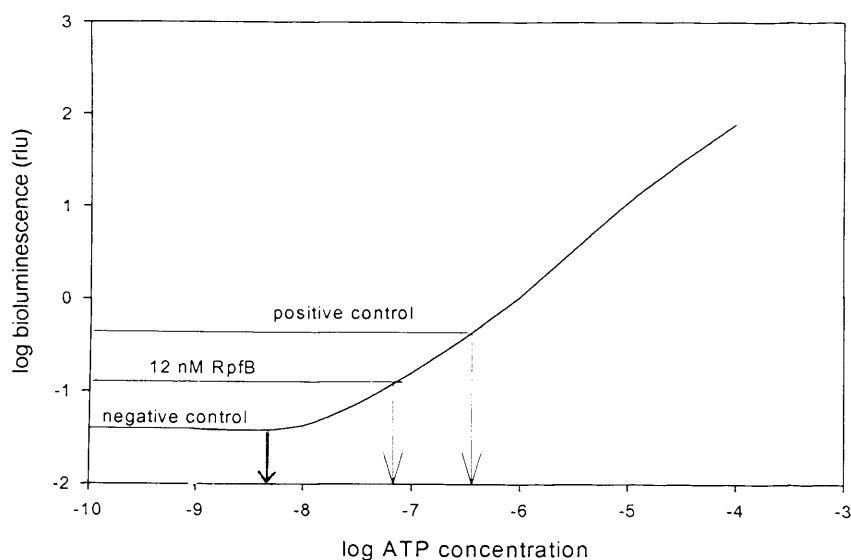
**Figure 5.6**

ATP bioluminescence of positive control (*M. luteus* supernatant) for overnight activity test after 16 h incubation. ATP concentration in the control =  $4 \times 10^{-6}$  M. ATP concentration in the supernatant =  $1.8 \times 10^{-5}$  M. Hence there is a 4.5-fold increase in ATP concentration in the positive control compared to the negative control.



**Figure 5.7**

ATP bioluminescence overnight activity test result using 12 nM fresh RpfB from clone CA1002. ATP concentration in negative control  $<10^{-8}$  M; ATP concentration in positive control =  $3.5 \times 10^{-7}$  M; ATP concentration with 12 nM RpfB =  $6.3 \times 10^{-8}$  M. Hence ATP concentration with 12 nM RpfB is  $>6$  x negative control and ATP concentration with positive control is  $>35$  x negative control.



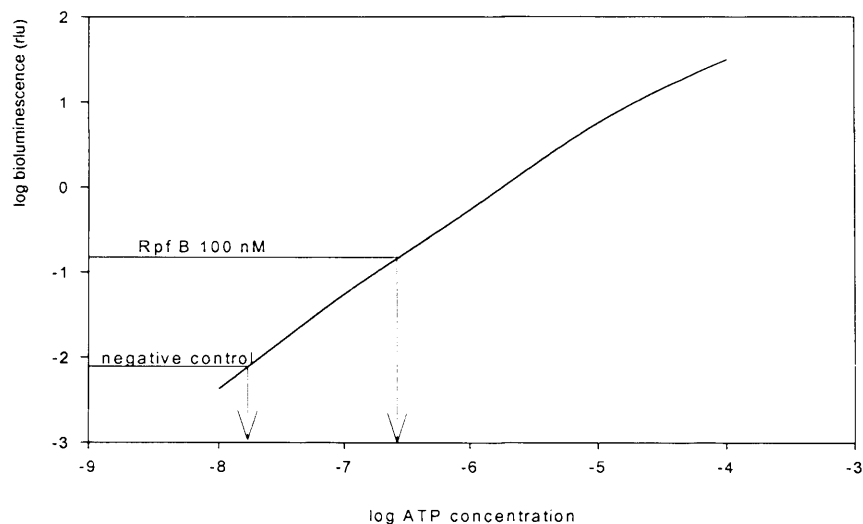
5.7.1.2.2      *Results: overnight ATP bioluminescence activity testing of RpfB from clone AD5 (pET19b vector and Origami B (DE3) pLysS host).*

Activity of RpfB from clone AD5 was measured three times. The protein was used immediately after production on each occasion. No protease inhibitors were used. In the first two experiments activity was greatest at RpfB concentrations of 60 - 100 nM, with a 5 and 12 x greater ATP concentration respectively in the RpfB culture compared with the negative control (see figures 5.8 and 5.9). In both cases, at concentrations of 6 - 10 nM there was a small increase in activity; at lower concentrations no effect was detected. Figure 5.10 demonstrates the measured activity of RpfB at different concentrations in one of these experiments. A bell-shaped curve is seen, in which the optimal activity is seen at an RpfB concentration of 60 nM, with less activity at concentrations of 600 nM and 6 nM and below. A similar pattern was seen in experiments with *M. luteus* Rpf on *M. luteus* cells (Mukamolova 1998) - except that the optimal activity was in the picomolar range - and with RpfA, C, and D on *M. luteus* and *M. smegmatis* (Mukamolova 2002b) - with optimal activity varying from subpicomolar - picomolar. Only RpfE has shown no evidence of reduced activity at elevated (micromolar) concentrations, when tested using *M. luteus*.

In a third experiment, ATP bioluminescence was measured at 4, 8, 12 and 16 hours with RpfB at a concentration of 80 nM. No activity of Rpf was measurable until the 12 hour reading, when there was a 3 x greater concentration of ATP with RpfB than in the negative control; at 16 hours this had risen to a 7 x greater concentration (figures 5.11 and 5.12).

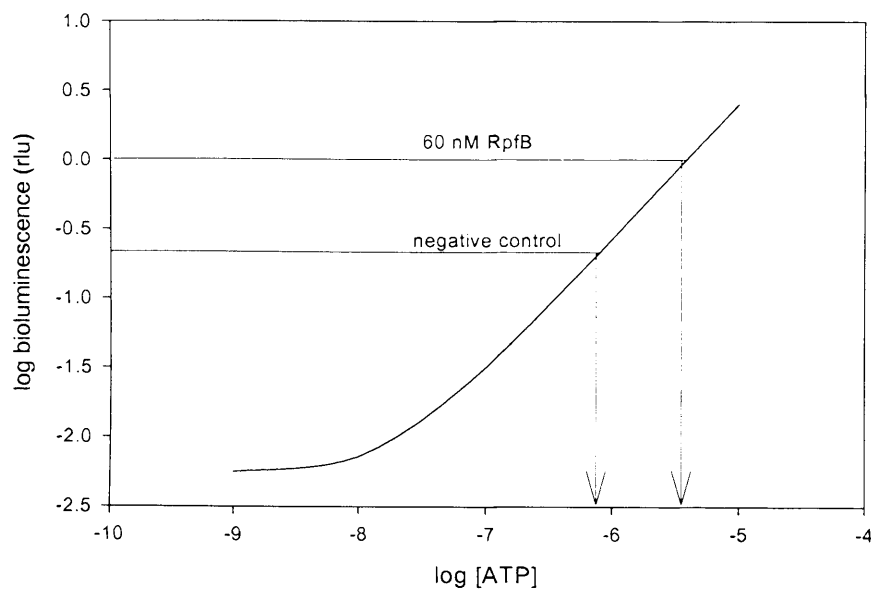
**Figure 5.8**

Result of overnight ATP bioluminescence activity test using 100 nM fresh RpfB from clone AD5. ATP concentration of negative control =  $1.5 \times 10^{-8}$  M; ATP concentration with RpfB at concentration of 100 nM =  $1.8 \times 10^{-7}$  M. Hence there is a 12 x greater ATP concentration with 100 nM RpfB.



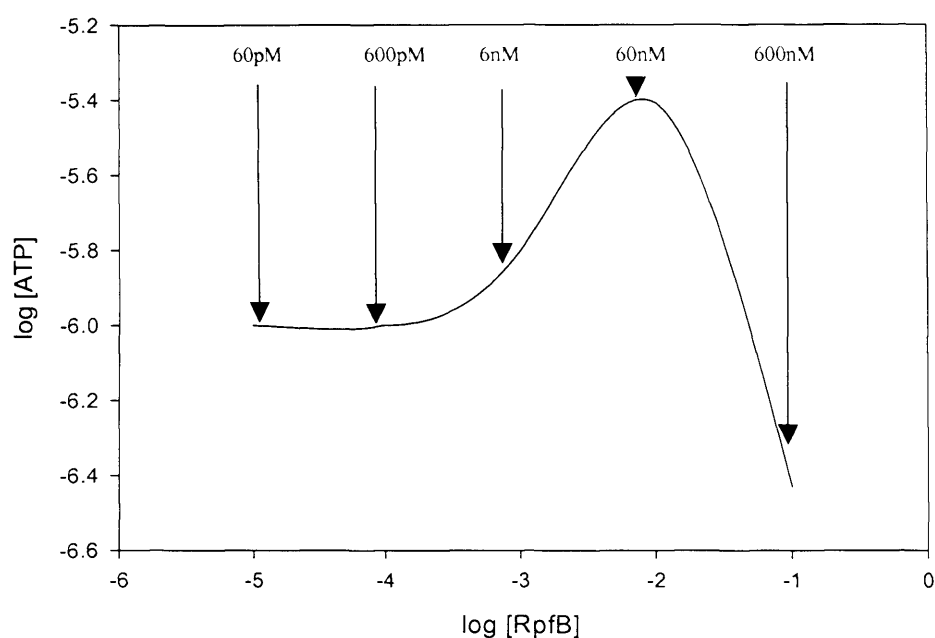
**Figure 5.9**

Result of ATP bioluminescence overnight activity test using 60 nM fresh RpfB from clone AD5. ATP concentration of negative control =  $7.5 \times 10^{-7}$  M; ATP concentration with 60 nM RpfB =  $3.8 \times 10^{-6}$  M. Hence sample containing 60 nM RpfB has 5 x greater ATP concentration.



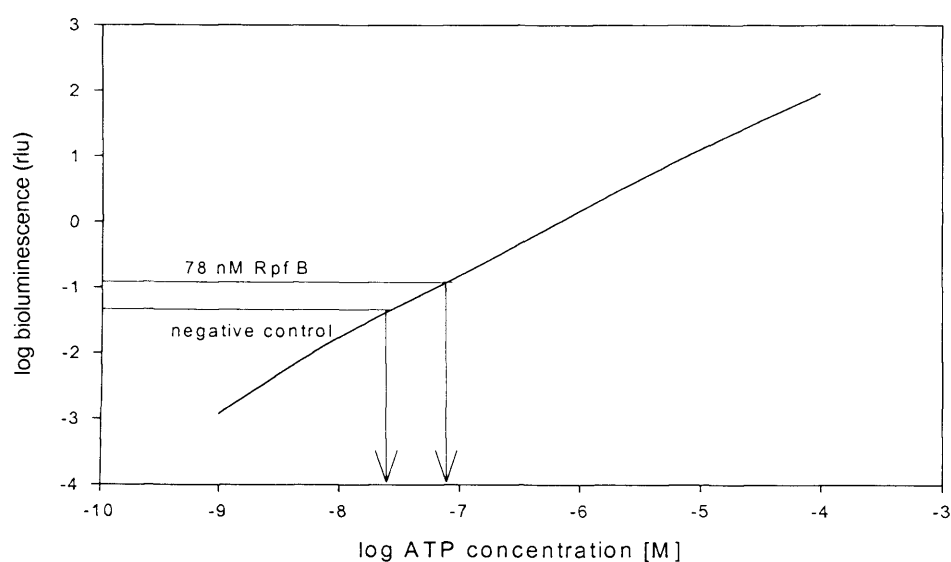
**Figure 5.10**

Activity of RpfB at different dilutions of 6.3  $\mu\text{M}$  stock solution, measured by ATP bioluminescence of *M. luteus* cultures. The greatest activity is seen at a concentration of 60 nM. Concentrations of 600 nM, 6 nM and less display less activity.



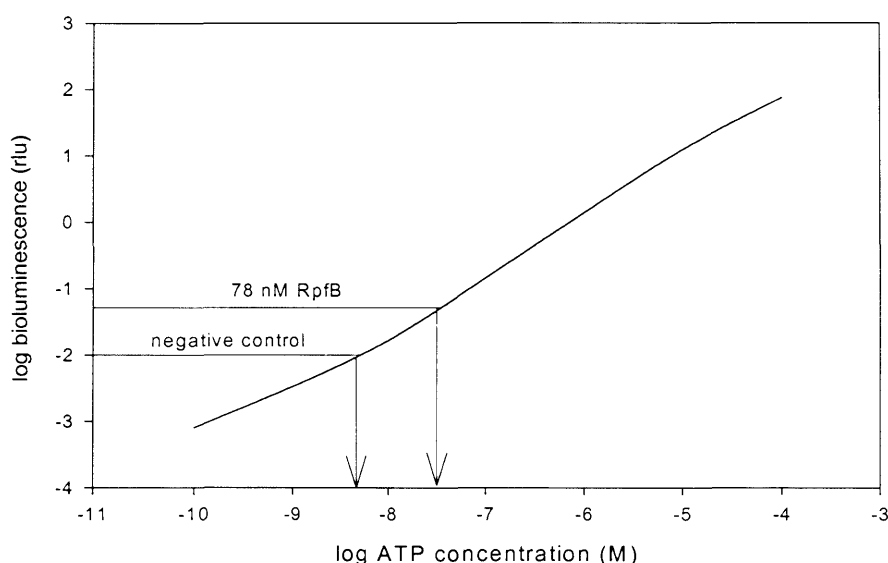
**Figure 5.11**

Results of ATP bioluminescence overnight activity test measured at 12 hours, using 80 nM RpfB from clone AD5. ATP concentration in negative control =  $2.2 \times 10^{-8} \text{ M}$ ; ATP concentration with 80 nM RpfB =  $7.1 \times 10^{-8} \text{ M}$ . Hence 3 x greater ATP concentration with RpfB detected at 12 h.



**Figure 5.12**

Results of ATP bioluminescence overnight activity test measured at 16 hours, using 80 nM RpfB from clone AD5. ATP concentration in negative control =  $3.98 \times 10^{-9}$  M; ATP concentration with 80 nM RpfB =  $2.8 \times 10^{-8}$  M. Hence 7 x greater ATP concentration with RpfB detected at 16 h (cf 3 x greater at 12 h, see figure 5.11).

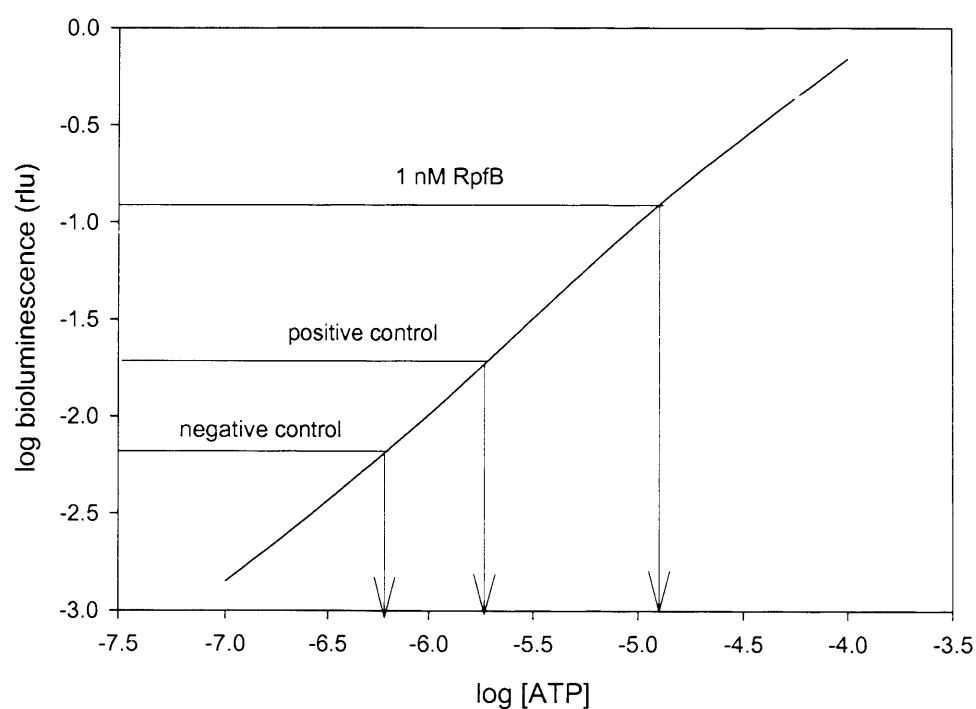


**5.7.1.2.3**      *Results: overnight ATP bioluminescence activity testing of RpfB from CA1002 (pET19b vector and HMS174 host)*

This was also measured on three occasions. Both with and without protease inhibitors, the most active concentration was found to be in the range 1-12 nM, which resulted in a 22-63 x greater ATP concentration in the cultures with RpfB compared to the negative controls. There was evidence of some activity in concentrations as low as 1 pM, but less so. One of these batches of protein was re-tested for activity after 6 days kept neat in the fridge. No activity could be then demonstrated. The results are shown in Figures 5.7 and 5.13.

**Figure 5.13**

Results of ATP bioluminescence overnight activity test using 1 nM concentration of 1 day-old RpfB from clone CA1002. ATP concentration of negative control =  $5.1 \times 10^{-7}$  M; ATP concentration of positive control =  $2.5 \times 10^{-6}$  M; ATP concentration with 1 nM RpfB =  $1.1 \times 10^{-5}$  M. Hence sample containing 1 nM RpfB has 22 x greater ATP concentration than negative control.



See Table 5.7 for a summary of the results of the overnight activity testing using ATP bioluminescence.

**Table 5.7**

Summary of results of overnight ATP bioluminescence activity tests with RpfB.

Clone used	Protease inhibitor used?	Age of protein	Stock concentration	Most active concentration	Increase in ATP cf negative control
AD5	no	fresh	6.3 $\mu$ M	60 nM	x 5
AD5	no	fresh	11.25 $\mu$ M	100 nM	x 12
AD5	yes	9 hrs	7.8 $\mu$ M	80 nM	x 7
CA1002	yes	24 hrs	10 $\mu$ M	1 nM	x 22
CA1002	yes	6 days	10 $\mu$ M	none	none
CA1002	no	fresh	12.6 $\mu$ M	12 nM	>x 6

#### 5.7.1.3 *Conclusions – overnight ATP bioluminescence activity test*

The overnight ATP bioluminescence activity test gave results within 24 hours which were discriminatory between batches, and was less laborious to set up than the methods described previously. The increase in ATP varied from x 5 - 22; the most active concentrations of RpfB were 1 - 100 nM (10  $\mu$ M - fM were tested). As previously found for Rpf, there is an optimal concentration for activity above and below which activity declines (figure 5.10). Difference in ATP was not detectable until 12 hours incubation at earliest and, of the times tested, was maximal at around 16 hours.

Recombinant RpfB from both the clones tested showed activity. If anything, RpfB from clone CA1002 (pET 19b vector, HMS174 host) showed a higher level of activity – the reverse was expected as clone AD5 was expressed using Origami B (DE3) pLysS host, which should, by facilitating disulphide bond formation in the cytoplasm, result in better-folded protein with potentially better biological activity. Use of protease inhibitors appeared to make little difference to the activity. Protein which was active

when first tested at 24 hours, was inactive when retested after 6 days storage in the refrigerator, though 24-hour old Rpf still had good activity.

### 5.7.2 *Rapid ATP bioluminescence activity test*

#### 5.7.2.1 *Method*

Having developed a successful assay for measuring Rpf activity using ATP bioluminescence, the next step was to develop a version of the assay which could give a result within one working day, so that active Rpf could then be used in further experiments as soon as possible. A similar overnight culture of *M. luteus* in nutrient broth was used as the starting inoculum, but only diluted about 20-fold in LMM rather than 1000-fold, to give an OD<sub>600</sub> in LMM of 0.1. The rest of the assay was performed in the same way as the overnight assay method (for details see Methods sections 2.5.3.1 and 2.5.3.2). To characterise the model, ATP in separate aliquots of culture was measured at different times up to 76 h. Early on in the experiment measurements were taken two-hourly; later they were taken at longer intervals. The aim was to use this model as a test to be read after a few hours rather than requiring an overnight incubation.

#### 5.7.2.2 *Results: rapid ATP bioluminescence activity test*

##### 5.7.2.2.1 *ATP levels over 76 hours*

Figure 5.14 shows an example of the results obtained using this model, and measuring ATP at different times up to 76 hours. ATP peaked more rapidly in the presence of RpfB - at 4 hours with RpfB and 10 hours without RpfB, reflecting a shorter lag phase with RpfB - then decreased. Also, ATP achieved a peak four times greater in the presence of RpfB. Further measurements of ATP concentration taken at time points up

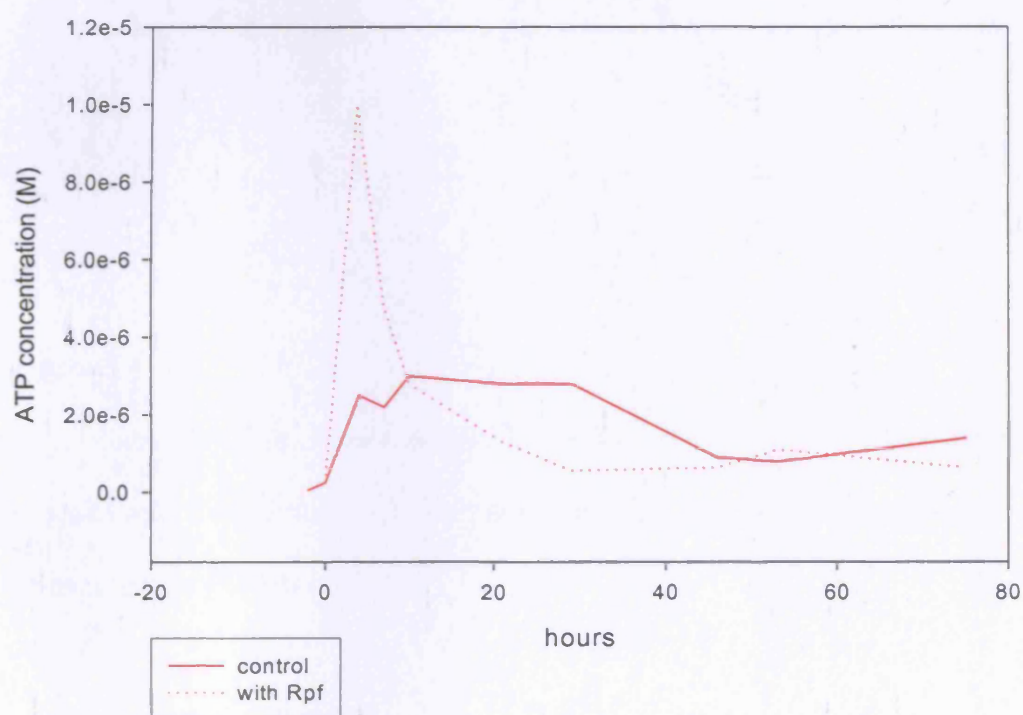


to 76 hours showed that the concentration of ATP then levelled off in both test and control samples, eventually reaching a similar level after 40 hours or so.

**Figure 5.14**

ATP concentration in *M. luteus* cultures in LMM with and without RpfB, over 76 h.

ATP peaked more rapidly in the presence of RpfB - at 4 h, compared with 10 h without RpfB, reflecting a shorter lag phase with RpfB. ATP achieved a peak 4 times greater in the presence of RpfB.



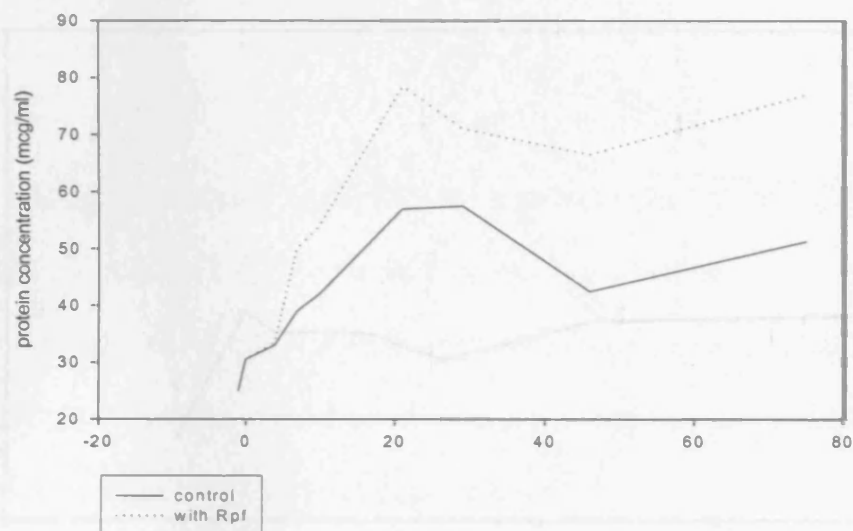
#### 5.7.2.2.2 Results: rapid ATP bioluminescence activity test – $OD_{600}$ and protein levels over 76 h

The  $OD_{600}$  and the protein content of the samples were measured to see how they correlated with the ATP concentration at intervals over periods of up to 76 hours. As might be expected, rises in  $OD_{600}$  and protein content lagged behind rises in ATP concentration: peaks in protein and  $OD_{600}$  were seen at ~12 - 24 hours of incubation, when organisms reached stationary phase. Higher peaks of both parameters were

obtained in the presence of Rpf, although the difference in both OD<sub>600</sub> and protein content reached only 1.5 - 2-fold, compared with the 4 - 5-fold difference seen in the ATP concentration. Figures and 5.15a and b (protein) and 5.16 (OD<sub>600</sub>) show examples of the results obtained.

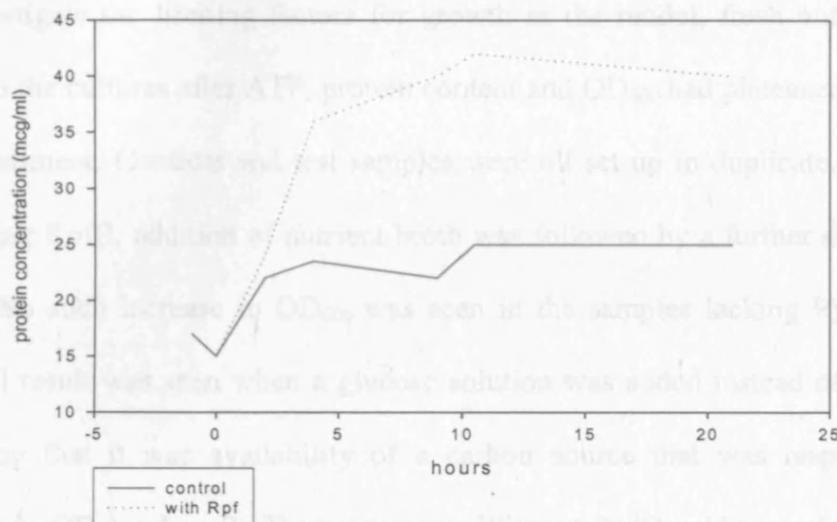
**Figure 5.15a**

Protein content of *M. luteus* culture in LMM with and without RpfB, measured at intervals over 76 h. The rise in protein content lagged behind the rise in ATP concentration (figure 5.14), peaking at ~20 h, and was 1.5-fold higher in the presence of RpfB.



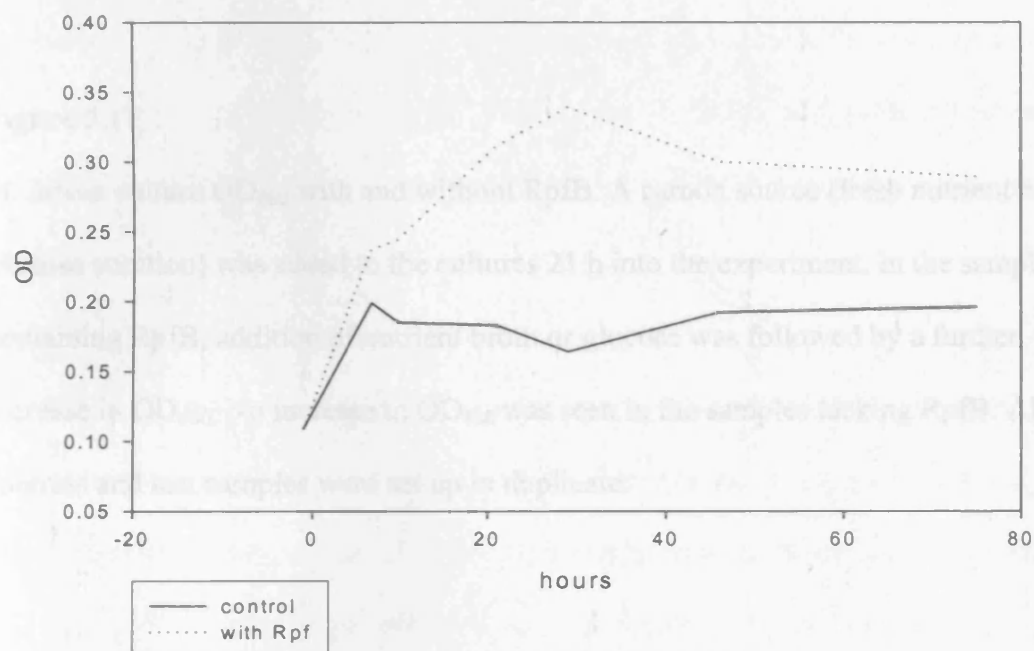
**Figure 5.15 b**

Protein content of *M. luteus* culture in LMM with and without RpfB, measured at intervals over 24 h. Protein content peaked at ~12 h and was approaching 2-fold higher in the presence of RpfB.



**Figure 5.16** Increase in OD<sub>600</sub>. These results are shown in figure 5.17, and indicate that

OD<sub>600</sub> of *M. luteus* cultures in LMM with and without RpfB, measured at intervals over 76 h. The rise in OD<sub>600</sub> approximately mirrored that of the protein content. Peak OD<sub>600</sub> was seen at ~12 - 24 hours of incubation, with a peak 1.5 - 2 -fold higher in the presence of RpfB.



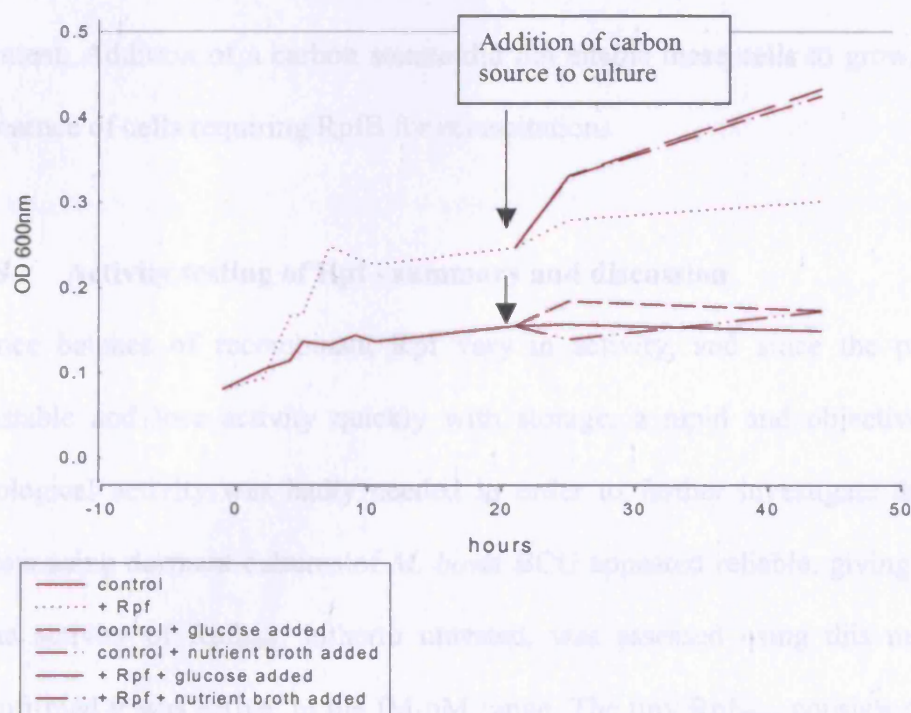
#### 5.7.2.2.3 Results: rapid ATP bioluminescence activity test – limiting factors for growth

To investigate the limiting factors for growth in the model, fresh nutrient broth was added to the cultures after ATP, protein content and OD<sub>600</sub> had plateaued, 21 hours into the experiment. Controls and test samples were all set up in duplicate. In the samples containing RpfB, addition of nutrient broth was followed by a further sharp increase in OD<sub>600</sub>. No such increase in OD<sub>600</sub> was seen in the samples lacking RpfB. An almost identical result was seen when a glucose solution was added instead of nutrient broth, indicating that it was availability of a carbon source that was responsible for the increase in OD<sub>600</sub> when RpfB was present. Without RpfB, addition of glucose did not

result in any increase in OD<sub>600</sub>. These results are shown in figure 5.17, and indicate that the growth of cells in the presence of RpfB is only limited by the availability of a carbon source in the minimal medium; however without Rpf, the cells still cannot grow even with the addition of the carbon source. Thus cells are present which cannot grow without Rpf. These cells may be considered to be in a dormant state and unable to leave this state in the absence of Rpf.

**Figure 5.17**

*M. luteus* culture OD<sub>600</sub> with and without RpfB. A carbon source (fresh nutrient broth or glucose solution) was added to the cultures 21 h into the experiment. In the samples containing RpfB, addition of nutrient broth or glucose was followed by a further increase in OD<sub>600</sub>. No increase in OD<sub>600</sub> was seen in the samples lacking RpfB. All controls and test samples were set up in duplicate.



#### 5.7.2.3 *Conclusions: rapid ATP bioluminescence activity test*

In this model, RpfB caused an early 4 - 5-fold increase in ATP content of cultures, which was detectable at 4 hours. The model was therefore found to be suitable for use as a rapid test of Rpf activity. However, the increase in ATP compared with the negative controls (without Rpf) was not as marked as it was in the overnight activity test, and so the level of discrimination between different batches was not as good. Increases in protein and OD<sub>600</sub> of 1.5 - 2-fold were also seen in the cultures containing RpfB, compared with the controls. These rises in OD<sub>600</sub> and protein content were detectable at 12-24 hours, lagging behind the rise in ATP concentration, as might be expected.

Growth of cultures containing RpfB was limited only by availability of a carbon source. Addition of nutrient broth or glucose solution allowed the cells to resume growth after it appeared to have plateaued at 21 hours. Cultures lacking RpfB appeared to contain cells that could not commence growth at all, as evidenced by a lower final OD<sub>600</sub> and protein content. Addition of a carbon source did not enable these cells to grow, reflecting the presence of cells requiring RpfB for resuscitation.

### 5.8 **Activity testing of Rpf - summary and discussion**

Since batches of recombinant Rpf vary in activity, and since the proteins appear unstable and lose activity quickly with storage, a rapid and objective test of their biological activity was badly needed in order to further investigate their properties. Tests using dormant cultures of *M. bovis* BCG appeared reliable, giving useful results. The activity of Rpf<sub>smeg</sub>, hitherto untested, was assessed using this method. Results confirmed it was active, in the fM-pM range. The tiny Rpf<sub>smeg</sub> consists only of what is believed to be the conserved active Rpf domain, and these experiments were the first



time that this domain alone, in a naturally-occurring, as opposed to artificially truncated form, had been tested for activity. The results confirm that this domain is the active region in the Rpf molecules.

Despite giving reliable results, the activity tests using *M. bovis* BCG took 2-3 weeks to give a result. Added to this was the fact that dormant cells of the organisms must be available, and these took 3-6 months to prepare.

The more rapidly growing *M. luteus* is also responsive to Rpfs, including those from *M. tuberculosis*. Therefore activity testing using this organism was developed instead. Tests relying on OD<sub>600</sub> using *M. luteus* in microtitre plates were too unreliable and unreplicable to be useful, because of the tendency of the organism to clump, combined with the very small inocula required to demonstrate the effect of Rpf. The approach devised to find a more rapid test of activity was to look at the concentration of ATP in cultures of stationary phase *M. luteus* inoculated at low density into a minimal medium. The amount of ATP increases in response to the presence of Rpf (as would be expected, since Rpf stimulates growth) and its measurement provided a suitable means to assay activity. Based on this, two models were developed – an overnight assay, and a rapid 4 hour assay using a higher starting inoculum.

In the overnight assay the increase in ATP varied from x 5 - x 22 in the presence of RpfB to > x 35 with the positive control. This assay demonstrated a more marked response to Rpf, with better discrimination between activities, than the rapid assay (a 5-fold difference at most). However difference in ATP concentration using the overnight assay method was not detectable until 12 hours' incubation at earliest. Thus its usefulness was somewhat limited by the fact that it required an overnight incubation, so

that by the time the result was known for a particular batch intended for use in further experiments, that batch could be expected to have lost some of its activity. The most active concentrations of RpfB in the range tested (10  $\mu$ M - fM ) were 1 - 100 nM. The range of activity of RpfB against *M. luteus* had not been measured previously. Most other Rpfs studied have had optimal activity in picomolar concentration (Mukamolova 2002b). Since RpfB, unlike the others tested, is thought to be membrane bound rather than secreted, it was plausible that its optimal concentration for activity might also differ. However, RpfB in picomolar concentrations was active against *M. bovis* BCG (Mukamolova 2002b). An alternative explanation for the nanomolar optimal activity found here could be the different method used to purify the protein in these experiments (using Novagen His-tag columns rather than in-house sepharose columns, with different buffers), which might in some way have affected the protein folding process or its stability.

In the rapid assay, RpfB in nanomolar concentrations caused a 4-5 -fold increase in ATP content of cultures containing RpfB compared with those without. The difference was detectable at 4 h – making it suitable as a rapid test of Rpf activity. However the increases in ATP levels with RpfB were not so marked as in the overnight assay, and showed less discrimination between batches. This test is useful for rapid confirmation that a batch of Rpf is active. Thus the choice of test will depend upon the purpose for which it is being carried out, with the overnight assay more appropriate for sensitive discrimination of activity level between batches, and the rapid assay as a screen to ensure active protein is present prior to using the protein in other experiments.



## **CHAPTER VI**

### **RESULTS: USE OF PHAGE DISPLAY TO IDENTIFY POSSIBLE CANDIDATES FOR AN Rpf RECEPTOR**

## 6.1 Introduction

The aim of the work described in this chapter was to see whether a possible receptor or receptors for Rpfs in *M. tuberculosis* could be identified. The Rpfs are strikingly potent cellular signals. This high potency, and the fact that they have an optimum concentration for activity above which they are inactive - a phenomenon seen in studies of ligand-receptor interactions - support the suggestion that they bind to specific receptors. One obvious way of gaining insight into the activity of the Rpfs is to identify these receptor(s). There are a number of ways of doing this. However many of the straightforward affinity techniques of receptor isolation are difficult to do with the mycobacteria due to their slow growth and their pathogenicity. An alternative method is to use a genomic technology linked to affinity isolation. The technique of choice is phage display.

### 6.1.1 Phage display

Phage display is a technique whereby heterologous proteins are expressed on the surface of phage particles. The phage become vehicles of expression which carry within them the genomic information for the expressed proteins, and which also have the capacity to replicate. Huge numbers of variant nucleotide sequences can be converted into populations of variant peptides and proteins which can be screened for desired properties. For example, phage display has been used to identify platelet-binding domains in fibrinogen-binding proteins of *Staphylococcus aureus* (Heilmann 2002) and fibronectin-binding proteins, IgG-binding polypeptides, adhesins, and transport proteins in the same organism (Jacobsson 1995, Rosander 2002). It is an extremely powerful tool for selecting proteins with specific binding properties from vast numbers of variants, and has been in use since 1985 (Smith 1985).

In order for a protein to be expressed at the phage surface, the nucleotide sequence encoding it is incorporated into a phage or phagemid genome as a fusion to a gene encoding a phage coat protein. As phage particles are assembled, the protein to be displayed is presented at the surface of the mature phage while the sequence encoding it is contained within the same phage particle. This physical link between the phenotype and genotype of the expressed protein, together with the replicative capacity of the phage, are the basis of phage display technology (Willats 2002). Libraries of nucleotide sequences of billions of fragments can be converted into populations of displayed variant proteins to be screened. Screening is usually done by an affinity selection (bio-panning) process during which phage populations are exposed to targets to selectively capture binding phage. Through successive rounds of binding selection followed by amplification of the selected phage, the original diverse phage population is enriched with phages which bind the desired target.

#### *6.1.1.1 Phage display vehicles*

Various different types of phage display system have been developed. The phage normally used are *E. coli* filamentous phage which infect their host by attaching to F-pili, and progeny phage are secreted without killing the host. Five structural proteins constitute the phage coat and they are inserted into the *E. coli* inner membrane before incorporation into the phage particle. Most commonly, the minor coat protein (pIII) or the major coat protein (pVIII) are used to display the foreign peptides (Jacobsson 2003). Different strategies have been employed to produce the fusion proteins. The gene fusion may be constructed between the nucleotide sequence and the coat protein gene in the viral genome. All copies of the chosen coat protein become fusion proteins – giving a high number of expressed foreign proteins, but with the risk of affecting phage viability if the functionality of the fusion coat protein is affected. Alternatively, hybrid phage

may be produced in which the gene fusion is an additional element of the phage genome, so that a wild-type copy of the gene is retained and phage express both wild-type and fusion proteins. Another way to produce hybrid phage is to use a phagemid based system (Willats 2002, Rosander 2002, Jacobsson 2003, Heilmann 2002). Sequences encoding fusion proteins are carried by plasmids with a phage origin of replication (phagemids). Most of the rest of the genes required for phage replication are carried by helper phage which are co-infected with the phagemids into host bacteria. In a hybrid system, fusion to pIII often results in less than one copy of the fusion protein per phage particle; fusion to pVIII gives multivalent display, the precise number of proteins per phage depending on factors such as the size, folding and amino acid composition of the foreign protein. Multivalent display may result in more efficient selection, but increases the risk of weak background interactions (Jacobsson 2003). In this work, a hybrid phage system using phagemid pG8H6 (supplied by B. Henderson) was used (Jacobsson 1996; derived from Jacobsson 1995), resulting in pVIII fusion proteins.

#### *6.1.1.2 Library amplification*

The simplest approach is to immobilise target molecules (in this case, recombinant *M. tuberculosis* Rpf) to a support and then expose it to suspensions containing the phage. Washing to remove non-binding phage is performed, then bound phage are eluted from the target. This can be done by using solutions of low pH, high salt, reducing agents or enzymatic cleavage. Following elution, the bound phage are then amplified by culture in *E. coli*.

Some expressed sequences may be incompatible with phage propagation, and others may be susceptible to proteolysis during propagation. Factors such as this impose

constraints upon the diversity of expressed proteins. Also, during amplification, phage with lesser avidity for the target but better growth characteristics may be preferentially amplified.

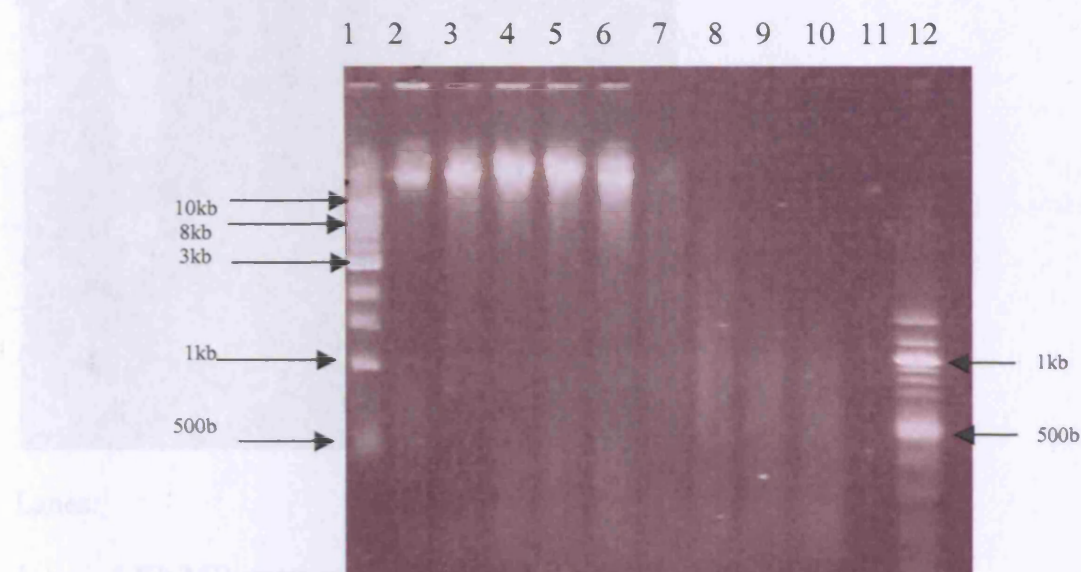
## **6.2 Construction of phage display library**

The method used is described fully in Methods section 2.6.1. Briefly, following isolation and purification, *M. tuberculosis* DNA was cut with blunt-ended restriction enzymes *Hae*III, *Rsa*I, *Alu*I, and *Hpy*CH4V. The aim was to find a concentration of DNA, restriction enzymes and lengths of digestion times which would provide a library of fragments between ~1.5 kb and ~200 bp in size. Figure 6.1 shows fragments of DNA of the required size range run on an agarose gel, after an appropriate protocol had been identified (as described in Methods 2.6.1). Figure 6.2 shows the result of scaling up the reaction tenfold. From this reaction, the 25 min, 35 min and 60 min fractions were pooled and then cleaned using the Qiagen gel extraction kit.

The DNA fragments were ligated into plasmid pG8H6, then electroporation reactions were performed using electrocompetent TG1 cells. Plasmid DNA was extracted from the resulting transformants and digested with *Hind*III and *Pst*I to check for the presence and size of inserts in the plasmid, which were found in about 50% of the transformants. Insert sizes ranged from ~180 bp to 1500 bp. Figures 6.3 a, b and c show examples of 20 clones which were tested when constructing the library, with 50% showing inserts of varying sizes.

Three electroporation reactions were performed and the resulting transformants pooled.

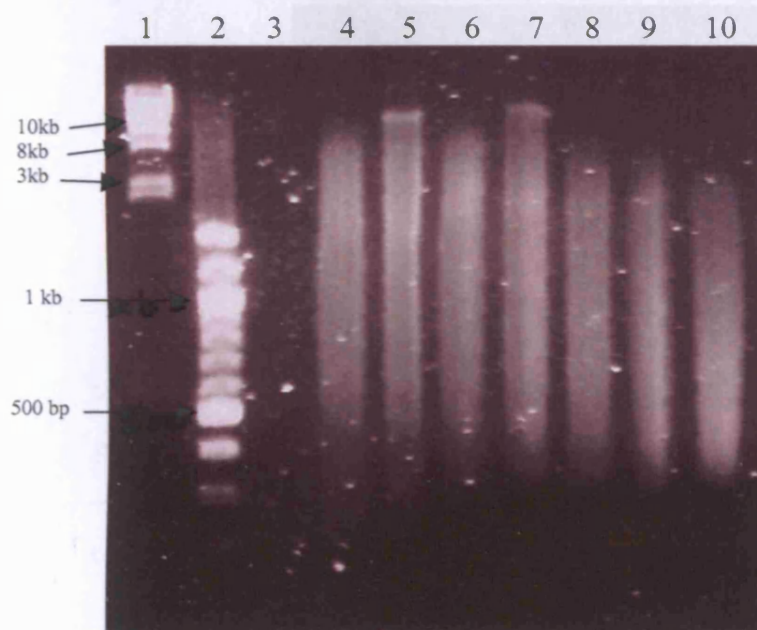
**Figure 6.1** Restriction enzyme cutting of *M. tuberculosis* DNA with blunt-ended restriction enzymes *Hae*III, *Rsa*I, *Alu*I, and *Hpy*CH4V. 5  $\mu$ l aliquots of digested DNA were removed at intervals of 2, 5, 15 and 35 min, heated to 80°C for 20 min to stop the reaction, and then run on a 1% TAE agarose gel. No cutting was apparent in the samples containing only 1  $\mu$ l of diluted enzyme, but a range of sizes of fragments was obtained, as desired, in the samples containing 2  $\mu$ l of diluted enzyme.



Lanes:

- |    |  |
|----|--|
| 1  | 1 Kb MW marker   |
| 2  | uncut <i>M. tuberculosis</i> genomic DNA                               |
| 3  | <i>M. tuberculosis</i> DNA cut with 1 $\mu$ l of diluted enzyme, 2 min |
| 4  | “ “ “ “ “ “ , 5 min  |
| 5  | “ “ “ “ “ “ , 15 min   |
| 6  | “ “ “ “ “ “ , 35 min   |
| 7  | <i>M. tuberculosis</i> DNA cut with 2 $\mu$ l of diluted enzyme, 2 min |
| 8  | “ “ “ “ “ “ , 5 min  |
| 9  | “ “ “ “ “ “ , 15 min   |
| 10 | “ “ “ “ “ “ , 35 min   |
| 11 | Blank well   |
| 12 | 100bp MW marker  |

**Figure 6.2** The reaction shown in 6.1 was scaled up. The reaction was stopped in 3  $\mu$ l aliquots at 2 min, 5 min, 10 min, 15 min, 25 min, 35 min and 60 min, then run on a 1% TAE agarose gel as shown. The 25 min, 35 min and 60 min fractions were pooled and cleaned for use in the library.

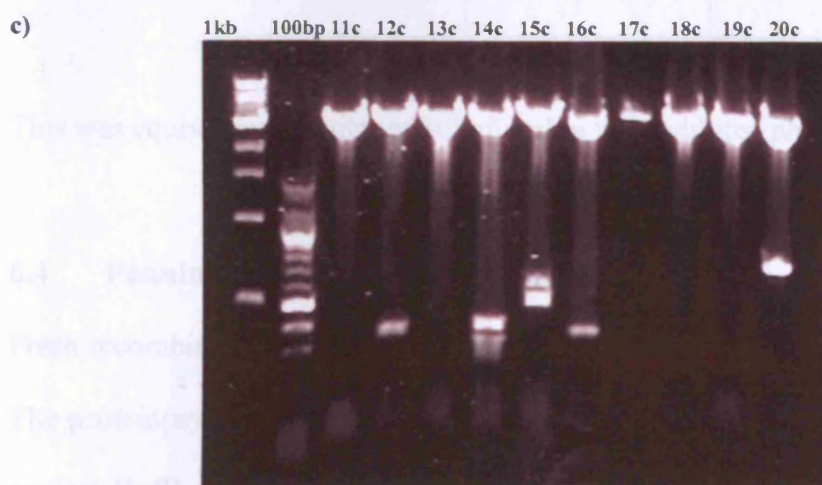
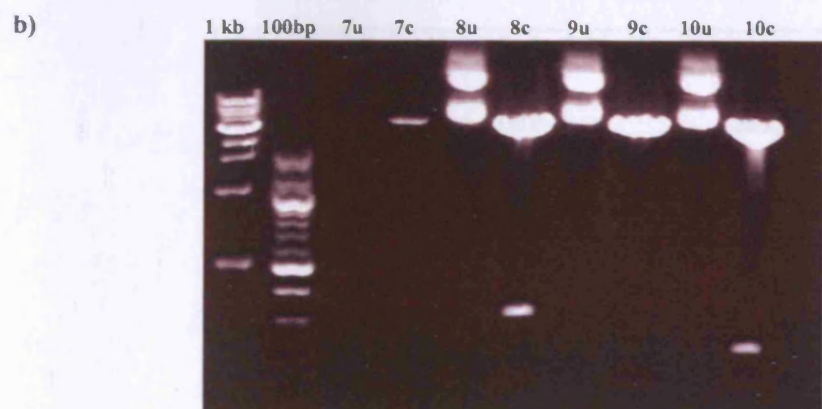
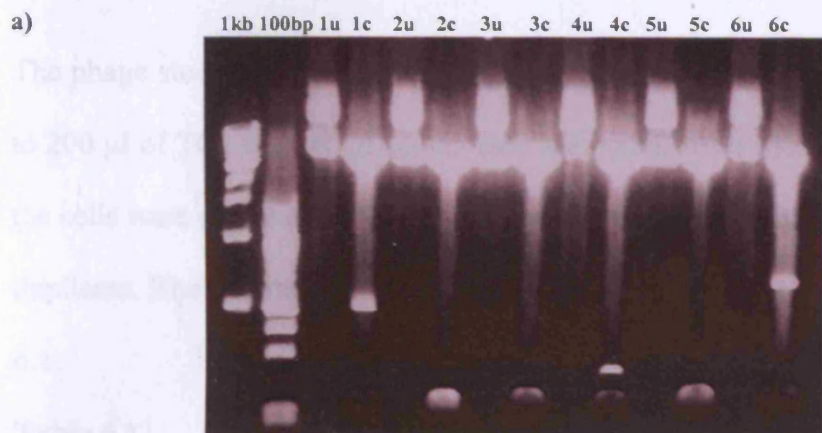


Lanes:

1	1 Kb MW marker
2	100 bp MW marker
3	blank lane
4	<i>M. tuberculosis</i> genomic DNA digest at 2 min
5	" " " " " 5 min
6	" " " " " 10 min
7	" " " " " 15min
8	" " " " " 25min
9	" " " " " 35min
10	" " " " " 60min



**Figure 6.3** Figures 6.3a, b and c show examples of the restricted DNA of 20 clones which were tested when constructing the library. Clones are numbered 1-20, with 'u' denoting uncut plasmid and 'c' denoting cut plasmid. 50% of the clones show inserts of various sizes.





### 6.3 Transfection of phage, PEG precipitation of phage and titering of phage stocks

The cells containing the phagemid library were transfected with helper phage R408, and the phage precipitated as described in the Methods section.

The phage stocks were titred by adding 10  $\mu$ l of  $10^{-5}$  -  $10^{-8}$  dilutions of the phage stock to 200  $\mu$ l of TG1 cells at OD<sub>600</sub> 0.5 (see Methods). After incubating for 30 min at 37°C the cells were plated out in 100  $\mu$ l aliquots on plates containing 100  $\mu$ g/ml ampicillin, in duplicate. The colonies were counted the next day and the results are shown in Table 6.1.

**Table 6.1**

Dilution of stock	cfu in 100 $\mu$ l	
	plate 1	plate 2
$10^{-8}$	1	1
$10^{-7}$	10	9
$10^{-6}$	84	70
$10^{-5}$	Semi-confluent	Semi-confluent

This was equivalent to a titre of  $10^9$  pfu/ml in the undiluted phage stock.

### 6.4 Panning for RpfB receptor protein

Fresh recombinant RpfB was used, produced as described in Methods section 2.2 – 2.3. The protein was dialysed in PBS to remove urea and imidazole. Panning of the library against RpfB was then performed as described in Methods 2.6.4. Fractions collected from panning were titred in duplicate (Methods 2.6.3) and plated out in 100  $\mu$ l volumes per plate.

#### 6.4.1 First round of panning

The phage titres were obtained in each fraction after the first round of panning are shown in Table 6.2.

**Table 6.2**

fraction	Dilution	cfu plate 1 (100 $\mu$ l )	cfu plate 2 (100 $\mu$ l )	Titre (pfu/ml)
unbound	$10^{-5}$	>1000	>1000	> $10^9$
	$10^{-6}$	200	320	
	$10^{-7}$	46	39	
	$10^{-8}$	23	13	
final wash	neat	132	136	$\sim 1.3 \times 10^3$
	$10^{-1}$	30	32	
	$10^{-2}$	17	18	
bound	neat	$\sim 1500$	$\sim 1500$	$\sim 10^5$
	$10^{-1}$	450	400	
	$10^{-2}$	105	87	
	$10^{-3}$	27	24	
	$10^{-4}$	8	-	

The results show that most phage did not bind to the RpfB-coated tube, but that a small proportion did bind and only came off the tube after the elution step.

The bound fraction of phage was amplified as described in Methods 2.6.2. After amplification, titering was carried out again. The results are shown in Table 6.3.

**Table 6.3**

Dilution of amplified bound phage from 1 <sup>st</sup> panning	cfu in 100 $\mu$ l	
	plate 1	plate 2
$10^{-8}$	57	19
$10^{-7}$	305	397
$10^{-6}$	Semi-confluent	Semi-confluent
$10^{-5}$	Semi-confluent	Semi-confluent

This was equivalent to approximately  $3 \times 10^9$  -  $3.8 \times 10^{10}$  pfu/ml and showed that the amplification process was working well.

#### 6.4.2 Second round of panning

After the second round of panning, only the bound fraction was titred, before it was amplified: see Table 6.4.

**Table 6.4**

Dilution of bound phage from 2 <sup>nd</sup> round of panning	cfu in 100 $\mu$ l	
	plate 1	plate 2
neat	~1800	~1800
$10^{-1}$	46	41
$10^{-2}$	0	6
$10^{-3}$	1	1
$10^{-4}$	0	0

This was equivalent to  $\sim 4.4 \times 10^3$  –  $1.8 \times 10^4$  pfu/ml. In a succesful panning experiment, a positive result is usually manifested by a 50 – 1000 fold increase in eluted phage

compared to the first panning (Jacobsson 2003). This is not seen here. However the cycle was continued through amplification to the third round of panning.

#### 6.4.3 Third round of panning

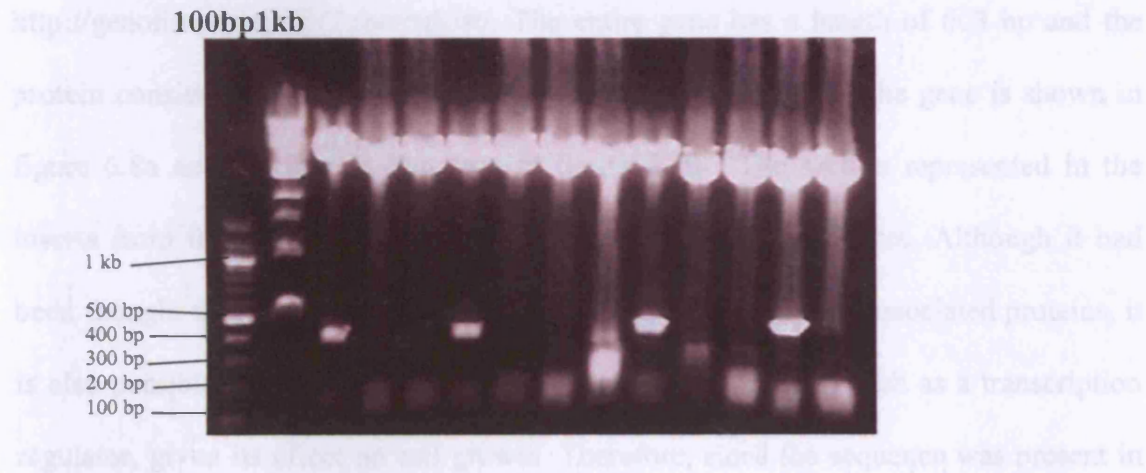
Counts on the bound fraction from the third round of panning are shown in Table 6.5.

**Table 6.5**

Dilution of bound phage from 3 <sup>rd</sup> round of panning	cfu in 100 $\mu$ l	
	plate 1	plate 2
neat	Semi-confluent	Semi-confluent
$10^{-1}$	530	389
$10^{-2}$	143	138
$10^{-3}$	35	13
$10^{-4}$	0	2

This represents a more satisfactory increase in the number of bound phage to approximately  $1 \times 10^5$ . Twelve colonies were picked to look for inserts. After overnight broth culture DNA was extracted and digested with *Pst*I and *Hind*III. The digests were examined by gel electrophoresis. Six of the twelve colonies contained inserts. All the inserts were below 500 bp in size and four of them had a similar size of about 400 bp (figure 6.4).

**Figure 6.4** 1% agarose gel electrophoresis of DNA from twelve clones obtained after three rounds of panning the phage library against RpfB. DNA was digested with *Pst*I and *Hind*III. Six of the twelve clones contained inserts of below 500 bp in size. Four of the six inserts had a similar size of about 400 bp.



## 6.5 Sequencing and analysis of inserts from clones selected by panning

To investigate the clones selected by the three rounds of panning, the inserts were sequenced, as described in 2.6.5. Three of the four inserts which had a similar size of about 400 bp had near-identical sequences; the other clones were different and contained no overlapping sequence with the three near-identical ones. The consensus sequences of the four inserts of size ~400 bp are shown in Figure 6.5, and an alignment of two of the near-identical sequences is seen in Figure 6.6. It was expected that the whole of each insert should correspond to a fragment of the *M. tuberculosis* genome (whether a coding sequence or otherwise). However this did not appear to be the case and each insert as a whole did not match with any part of the *M. tuberculosis* genome sequence, recorded at <http://genolist.pasteur.fr/TubercuList/>. Shorter fragments within the inserts did correspond to short fragments of the *M. tuberculosis* genome. The longest and most notable, highlighted in red in Figures 6.5 and 6.6, was a 17 bp sequence present in the three identical inserts, and identified by BLAST searches using the whole inserts against the genome, with an E value of 0.046.

This 17 bp fragment forms part of gene *Rv0328*, described as coding for a possible transcription regulatory protein, perhaps of the tetR/acrR family. Its co-ordinates within the *M. tuberculosis* genome are 394111 - 394713 and figure 6.7 is a diagram of its position within the surrounding genome region, taken from <http://genolist.pasteur.fr/TubercuList/>. The entire gene has a length of 603 bp and the protein consists of 200 amino acids. The nucleotide sequence of the gene is shown in figure 6.8a and the protein sequence in figure 6.8b. The section represented in the inserts from the panning experiments is highlighted in both figures. Although it had been thought that the Rpf receptor(s) were likely to be membrane-associated proteins, it is also possible that Rpf might bind inside the cell to a protein such as a transcription regulator, given its effect on cell growth. Therefore, since the sequence was present in half of the inserts selected during panning it was decided to clone the gene and express the protein, with a view to demonstrating whether or not it did bind Rpf.



**Figure 6.5** Consensus sequences of four sequenced inserts from clones isolated after three rounds of panning. The inserts in clones *b*, *c*, and *d* are near-identical. Green highlighted bases are those corresponding to amplified plasmid DNA; red highlighted bases are those found in three inserts to be present in *M. tuberculosis* gene Rv0328. The *Sma*I digest site which was used to cut the plasmid is underlined at both ends of the insert.

*Clone a*

CGGCCATGGCCAGGTGCAGCTGCAGCACCACCACCACCACCCCCC GCGCTGCACGCC  
GCCATGATCGCCGGACCTCGCCAACCGCCCGCGGAGGCGGAGCAGCTCGCCCGGATCGCC  
CTGGCCGAGACGCCGGAGCTGAACCTGCGGCTGGTGCAGATCGTCGCCGGCATCCTGGAGC  
GCGGCGGCAAGTCGGTGGAGGCGATGAAGCTGGTGGATAGCGTCACCGCCAGCGGGCGACGA  
CCTCTCCCTCTTCAGCACGCCCGAGGCGCGCCGGCGCCTGTTGCAGGGCCGCGCCGTCGCCT  
CGGCGGTCGAGGGCATGGCCGAGGCGCAG GGGTGCAC

*Clone b*

GGCCCAGGTGCAGCTGCAGCACCACCACCACCACCACCCCCC GGACAGGCGAAATCCGCTC  
CACGGCGCCGCCTCCCCACCCCCACGGGCCCGGGCGCGGCACAAAAAACGCCGGGGAAG  
CCTGCTTCCCCGGCGTGTCCCTGGCGGCAGGCGGGCAGTCTCCGCCCTGCCCTGCCCTGCCCT  
GGCCGGGCCTTACTCGGCCGGCTGCACCATGGCGGGGGCCTCGGC CGGGAGCGCCTGGCCG  
CT CAGCGCCCGGTTCGAG GGGTGCAGGTTCGACCTCGAGATCAAACGGGCGGCCCGCAGAAC

*Clone c*

GGACAGGCGAAATCCGCTCCACGGCGCCGCCTCCCCACCCCCACGGGCCCGGGCGCGGCA  
CAAAAAAACGCCGGGGAAGCCTGCTTCCCCGGCGTGTCCCTGGCGGCAGGCGGGCAGTCTC  
CGCCCTGCCCTGCCCTGCCCTGGCCGGGCCTTACTCGGCCGGCTGCACCATGGCGGGGGCCT  
CGGC CGGGAGCGCCTGGCCGC CAGCGCCCGGTTCGAG GGGTGCAGGTTCGACCTCGAGATCA  
AAC

*Clone d*

GGYCCAKGTGCAGCTGCAGCACCACCACCACCACCACCCCCC GGACAGGCGAAATCCGCTC  
CACGGCGCCGCCTCCCCACCCCCACGGGCCCGGGCGCGGCACAAAAAACGCCGGGGAAG  
CCTGCTTCCCCGGCGTGTCCCTGGCGGCAGGCGGGCAGTCTCCGCCCTGCCCTGCCCTGCCCT  
GGCCGGGCCTTACTCGGCCGGCTGCACCATGGCGGGGGCCTCGGCCGGGGC CGGGAGCGCC  
TGGCCGCT CAGCGCCCGGTTCGAG GGGTGCAGGTTCGACCTCGAGATCAAACGGGCGGCCGCA  
G

**Figure 6.6** Sequence alignment of two near-identical inserts from clones isolated after three rounds of panning. A 17 bp segment corresponding to a segment in *M tuberculosis* gene Rv0328 is highlighted in red.

Sequence 1	lcl seq_1	Length	227	(1 .. 227)
Sequence 2	lcl seq_2	Length	221	(1 .. 221)

```

Query:
ggacaggcgaaatccgctccacggcgccgcctccccacccccacggggcccgggcgcggc 60
|||||
Sbjct:
ggacaggcgaaatccgctccacggcgccgcctccccacccccacggggcccgggcgcggc 60

Query:
acnnnnnnncgccccgaagcctgcttccccggcgtgtccctggcggcaggcgggcagtc 120
||
Sbjct:
acaaaaaacgccccgaagcctgcttccccggcgtgtccctggcggcaggcgggcagtc 120

Query:
tccgccctgccctgccctgccctggccgggccttactcggccgggtgcaccatggcgggg 180
|||||
Sbjct:
tccgccctgccctgccctgccctggccgggccttactcggccgggtgcaccatggcgggg 180

Query: 181 gcctcggccggggc cgggagcgctggccgctcagcgcccggtcgag 227
|||
Sbjct: 181 gcctc-----ggc cgggagcgctggccgctcagcgcccggtcgag 221

```

Score = 41.1 bits (21), Expect = 1.1  
Identities = 25/27 (92%)





**Figure 6.8 a** Gene sequence of *M. tuberculosis* Rv0328

```
>M. tuberculosis H37Rv|Rv0328|Rv0328: 603 bp - POSSIBLE
TRANSCRIPTIONAL REGULATORY PROTEIN (POSSIBLY TETR/ACRR-FAMILY)
caagg
agaacacgtcgccgtacttgcgatgcatcgctggcagga
accattcccgaaacctcaggtacagcacgctctgcaggt
agcggggtagccgcgggcccgggtggcaggcccgcgtcga
acgtgcttgccatggcgggtcccttctgataatcaaata
tttgatgtaaacaatgcttatcacgataggatgcagct
1 - gtg caa cag caa cgc aca aac cgc gac aaa
31 - ctg ctc gac ggc gct ctg gct tgt tta cga
61 - gaa cgc ggc tac ggc aac acc agc tcg cgc
91 - gac atc gct cgt gcg gca ggg gtg aac atc
121 - gcg tcg atc aac tac cac ttc ggt agc aag
151 - gac gcg ctg ctc gac gat gcg ctc ggc cgg
181 - tgc ttt tcg acg tgg aac cag cgt gtc cag
211 - gag gca ttc gat cac tcc cgc gcc gcc ggt
241 - ccg gcc ggg cag atc ctg gcg gta ctc gaa
271 - gcc acc gtc gat tcg ttc gag cag atc cgc
301 - ccc gcc gtg tat gcg tgt gtg gag tca tac
331 - gct ccg gcg ttg cgc tca gag gcc ttg cgg
361 - gag cgc ctg gcc gcc gga tat gcc gac gtt
391 - cgg cag cat tcg gtc gat ctg gct ggc gct
421 - gcg ctt gcc ggt acc gac ata gca ccg ccg
451 - gag aac ctg tcg acc atc gtc tcg gtg ttg
481 - atg gcg gtc atc gat ggc ctc atg atc cag
511 - tgg atc gcc gat ccg tcc gcc acc ccg cga
541 - tcg acc gag gta atc cga gcg ctt gcc agc
571 - atc ggc gcg gtc gtc acg tcg cag ttg cgg
601 - tga
accacacgggtcgccggatggtctgcactgcgcttgatgc
cgacgtcgatgaagccggcagcgccaagccacgcggcgg
tgtcgagggtggggggcacgcgatagatcgcgggatcga
aacgggcccgcagtggttggtcatcggatatcgatgtaa
ggacgagtcgacccccgggcccgcagggtcgagcgatgt
cgcaa
```

**Figure 6.8 b** Protein sequence of *M. tuberculosis* gene Rv0328

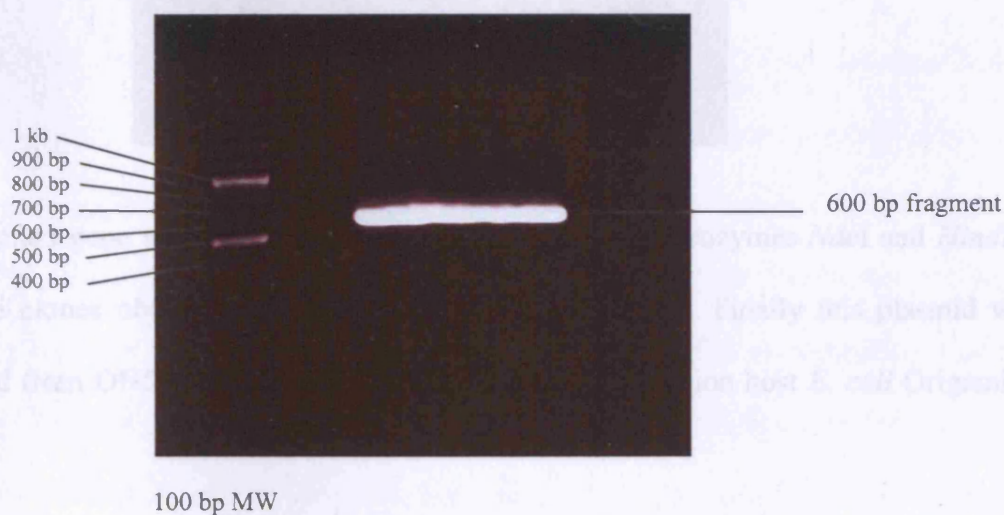
```
>M. tuberculosis H37Rv|Rv0328|Rv0328: 200 aa - POSSIBLE
TRANSCRIPTIONAL REGULATORY PROTEIN (POSSIBLY TETR/ACRR-FAMILY)
1 - VQQQRTNRDK LLDGALACLR ERGYNTSSR DIARAAGVNI ASINYHFGSK
DALLDDALGR
61 - CFSTWNQRVQ EAFDHSRAAG PAGQILAVLE ATVDSFEQIR PAVYACVESY
APALRSEALR
121 - ERLAAGYADV RQHSVDLAGA ALAGTDIAPP ENLSTIVSVL MAVIDGLMIQ
WIADPSATPR
181 - STEVIRALAS IGAVVTSQLR
```



## 6.6 Cloning and expression of Rv0328

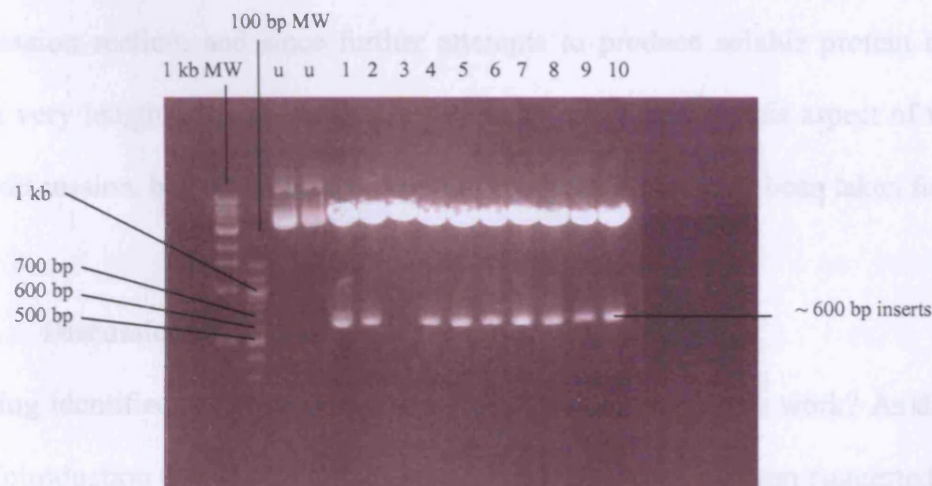
Primers were designed to amplify Rv0328 from *M. tuberculosis* genomic DNA. Primers had an *Nde*I site at the 5' end (primer sequence ATA TCA TAT GCA ACA GCA ACG CAC A) and a *Hind*III site at the 3' end (primer sequence ATA TAA GCT TCA CCG CAA CTG CGA CGT). The PCR reaction was carried out as described in Methods 2.1.4, and a ~600bp segment of genome was successfully amplified (Figure 6.9).

**Figure 6.9** Agarose gel showing amplification of Rv0328 from *M. tuberculosis* genomic DNA. A ~600bp segment of genome was successfully amplified. 50 µl of the reaction mixture was run on this gel, for subsequent excision and cleaning.



This 600 bp fragment was cloned into TOPO® 2.1 (2.1.5) and transformed into competent *E. coli* DH5α cells (2.1.11 and 2.1.12). After plating onto XGal/IPTG/Ap100 plates (2.1.13) for blue-white selection, approximately 3 white colonies were obtained per 100 µl of transformants plated (and no blue colonies). Of ten such clones subcultured and examined, nine contained the required 600 bp insert (figure 6.10).

**Figure 6.10** Inserts of cloned Rv0328 from clones of TOPO® 2.1 vector transformed into *E. coli* DH5 $\alpha$ , after blue-white selection. Of ten clones subcultured and examined, nine contained the required 600 bp insert. u = uncut plasmids; lanes 1-10 correspond to clones tested, only clone 3 lacks the required insert.



The cloned gene was then sub-cloned into pET-21a, using enzymes *Nde*I and *Hind*III. All the clones obtained again contained the 600 bp insert. Finally this plasmid was isolated from DH5 $\alpha$ , and transformed again into the expression host *E. coli* Origami B (DE3).

Protein expression was performed under the same conditions as described for the *M. tuberculosis* Rpf5 in 2.2.1 and 2.2.2, with extraction and purification as in 2.3.1 and 2.3.2.2, i.e. using sonication to break up the cells and Quick 900 nickel columns (Novagen) to purify the His-tagged protein. In addition, lysozyme was added to the cell suspension and left for half an hour prior to sonication, to try and improve the extraction. Extraction and purification was first attempted without the use of urea. The yield of purified protein in the eluted fraction was low at ~40  $\mu$ g/ml. SDS-PAGE showed that although a band was visible corresponding to the purified protein, much of

the protein remained insoluble in the cell deposit. The extraction was therefore repeated using 6M urea: this resulted in a small improvement in solubility (a yield of ~ 60 µg/ml in the eluted fraction). Whether or not this low concentration was sufficient to proceed to Rpf-binding experiments was not certain. The logical next step would have been to enhance the solubility of Rv0328 before performing binding experiments. However since there was doubt over the relevance of the interaction, as highlighted in the discussion section, and since further attempts to produce soluble protein might have been very lengthy, it was decided not to invest more time in this aspect of the project. See discussion, below, for how this part of the work might have been taken further.

## **6.7 Discussion**

Having identified the Rpfs an obvious question is how do they work? As described in the Introduction section 1.4.2 two possible mechanisms have been suggested. The most obvious is that the secreted or cell surface Rpfs act in an autocrine manner and bind to a very high affinity receptor on the cell surface. There are many examples in Gram-positive bacteria of cell-cell communication using modified oligopeptides that are secreted into the medium and accumulate at high cell density. The detectors for these oligopeptide signals are two-component adaptive response proteins, and the mechanism of signal transduction is via a conserved phosphorylation/dephosphorylation mechanism. A precursor peptide is synthesised, then modified to make the mature autoinducer molecule. This is secreted by an ATP-binding cassette (ABC) transporter complex. At a critical concentration of the oligopeptide in the medium, the two-component sensory recognition apparatus detects the signal. Sensory information is relayed into the cell by phosphorylation and culminates in an appropriate alteration in gene expression. The detector for the oligopeptide autoinducer is called the sensor kinase, and it transfers the phosphoryl signal to a downstream regulatory component,

the response regulator. This is a DNA binding protein whose function is usually to induce the expression of density-controlled target genes. Such systems are seen for example in *Bacillus subtilis*, to regulate sporulation, and in *Staphylococcus aureus*, regulating several pathogenicity determinants. Therefore there is a clear precedent for this type of system to be used for the RpfS. An alternative explanation is that the RpfS act as enzymes and either by cleaving the cell wall, or by cleaving the cell wall and releasing the key mediator, bacteria enter the cell cycle. It was the first hypothesis, considered the most obvious, that was being tested here.

Panning of the phage library against recombinant RpfB identified an identical 400 bp fragment in one in two of the resulting clones. This fragment contained a 17 bp sequence which matches a sequence in *M. tuberculosis* gene Rv0328, predicted to code for a transcription regulatory protein. This gene was successfully cloned but unfortunately the expressed protein had low solubility with the expression system used. Further efforts to express this protein were not pursued: the fact that only a 17bp matching sequence was detected, and that the rest of the 400bp fragment did not appear to match parts of the *M. tuberculosis* genome as it ought, placed the relevance of the interaction in some doubt. However, the possibility has not been excluded that the protein identified, Rv0328, does bind Rpf. The next step would have been to improve the solubility of the recombinant Rv0328 by using a different expression system. For example, a vector enabling a soluble fusion protein to be expressed, such as pET43.1, could have been used, as was done in Chapter 4 to improve the solubility of Rpf<sub>smeg</sub>. Once enough soluble protein was available, binding assays with Rpf could have been performed. One simple way to demonstrate binding would be to pass the His-tagged Rv0328 down a nickel column, followed by recombinant Rpf or supernatant of dormant *Micrococcus luteus* cultures. After saving the unbound fraction, the bound fraction

could be eluted from the column with imidazole, and subjected to SDS-PAGE, with the unbound fraction and Rv0328 run as controls. If binding had occurred, two bands should be seen in the test lane, one corresponding to Rv0328 and the other to Rpf.

Had such experiments failed to show binding of Rpf to the product of Rv0328, other possible reasons why a receptor for RpfB might not have been identified are summarized below. Firstly, the phage display library may not have been extensive enough. In a successful panning experiment, a number of clones with overlapping inserts are desirable, and not all variations of a protein may be displayed equally well. Therefore ideally the size of the library should exceed the 'minimum' number of clones several times, and perhaps the library used here was not large enough. Also, the method used to generate the DNA fragments, blunt-ended restriction, does not give truly random cleavage of the DNA, and perhaps sonication of the genomic DNA may have given a better result (Jacobsson 2003). The insert frequency in the library was at an acceptable level (phagemid with insert was found in 50% of colonies growing on ampicillin plates), but higher levels approaching 100% can be achieved using phagemids such as pG8SAET. In pG8SAET insertion of foreign DNA into a cloning site is required for expression of the fusion protein, as without the insertion gene VIII is out of frame with the signal sequence. However, numerous attempts to ligate *M. tuberculosis* DNA into pG8SAET were unsuccessful, for reasons which could not be identified. Therefore the library made with pG8H6 was used. It would have been informative to test the library to see whether it could identify a known receptor protein such as the fibronectin-binding proteins of the Antigen85 complex (Abou-Zeid *et al* 1991) and thus demonstrate that it was indeed functional as a library.

If the library was adequate, then it is a possibility that the Rpf was not binding well to the polystyrene tube – since one of the problems with recombinant Rpf is that it is insoluble, the protein may have been aggregating during dialysis to remove urea and imidazole. Passive adsorption to the tube relies on establishing a large number of relatively weak bonds between target and support which can result in the target being forced out of its functional configuration (Wilson 2001). Alternatively, the production of the recombinant protein in the presence of urea could mean that it was not folded in its natural tertiary conformation and so did not bind the receptor as it should.

It may also be the case that selection of clones was for some property other than specific ligand binding, such as replication or packaging into the phage particle (Jacobsson 2003). During amplification, phage with inferior avidities for the target but better growth characteristics may be preferentially amplified (Willats 2002) and there is evidence that this may have occurred in the experiments described here. According to Jacobsson's experience, a 'true' selection should be apparent where clones are found with different (overlapping) inserts derived from the same gene. Where only identical inserts are found (as is the case here), it suggests selection for some other property than ligand binding.

Finally, although it was originally supposed that Rpf's have a protein receptor, this may not necessarily be the case. The similarity of one of its folds to the active site of lysozyme (Cohen-Gonsaud 2004 and 2005) and further bioinformatic analysis (Ravagnani, Finan & Young 2005) has raised the possibility that its mode of action may be cleavage of an oligosaccharide in the cell wall, not involving a protein two-component response system at all. The evidence for this, although at first thought a less



likely mechanism, is now mounting, and is discussed in the Introduction and Discussion chapters.

## **CHAPTER VII**

### **RESULTS: IMMUNOCYTOCHEMISTRY USING ANTI-Rpf ANTIBODIES**

## 7.1 Introduction

### 7.1.1 Aim

The question of whether Rpf is expressed in *M. tuberculosis*-infected human tissues has not previously been addressed although Rpf expression has been shown in mice (Tufariello *et al* 2004). The aim of the work presented in this chapter was to investigate for the first time, by immunocytochemistry using anti-Rpf antibodies, whether the presence of Rpf could be detected in *M. tuberculosis*-infected human tissue sections, and if so, what its distribution was. This would provide the first demonstration that the organisms in human infection produce Rpf. The correlation of Rpf with ZN-stained bacilli in the sections was also investigated.

### 7.1.2 Histobacteriology of tuberculosis

In order to make sense of immunocytochemistry aiming to detect proteins produced by *M. tuberculosis*, an understanding of the histobacteriology of tuberculosis is required. An authoritative text on this subject, *Le bacille de Koch dans la lesion tuberculeuse du poumon*, by Georges Canetti, then director of the Pasteur Institute in Paris, was published in 1946, with an English language revised edition following in 1955 (Canetti, 1955). Canetti studied sections of human lung from 1500 autopsies and correlated the results with clinical and radiological data. He counted the numbers of acid-fast bacilli seen in the various types of histologic lesions, to try and get a clearer understanding of the evolution of each lesion. He also drew on Huebschmann's previous study of histogenesis of tuberculous lesions (Huebschmann 1929).

Classically, three main categories of tuberculous lesion were considered to exist. Firstly, an inflammatory or 'exudative' lesion, consisting of polymorphs, lymphocytes and macrophages; secondly, a 'productive' lesion, in which cellular metaplasias occur, with

macrophages and histiocytes forming epithelioid cells and then giant cells. The epithelioid cells are often arranged in a rounded formation known as a tubercle. Thirdly, the caseous lesion, in which tissue necrosis, followed by homogenisation, is seen. The homogenised centre may soften.

According to both Huebschmann and Canetti, the majority of human pulmonary tuberculous lesions start with an exudative-type reaction which is eventually followed by caseation. Caseation is followed by the 'productive' reaction with tubercle formation and sclerosis at the periphery of the caseum (this contradicted earlier dogma which had the tuberculous process starting as either exudative or productive, with both types then progressing to caseation). The evolution of a lesion is as follows: in the pre-exudative and exudative phases, the alveoli fill with fibrin, desquamated and swollen alveolar cells and white cells. The bundles of fibrin may be large. The exudative lesions may be sub-divided into fibrino-macrophagic, purely fibrinous or polymorphonuclear, depending on the proportions of fibrin and white cells present. This appearance is not pathognomonic of tuberculous infection – acid-fast bacilli, or small areas of caseation, must be seen to make a firm diagnosis. In the next stage, caseation begins. The fibrin loses its fibrillary structure and swells to fill the alveolus. The alveolar septa lose capillaries and become necrotic. The alveolar wall disintegrates so that several alveoli are amalgamated into blocks of necrosis. These may remain solid, and this is what occurs in latent infection, and in most active infections. However the caseum may soften and this is what leads to clinical tuberculosis: the liquid caseum is expelled via a bronchial segment, infecting adjacent areas of lung and leaving a cavity which is difficult to heal. Hence it is the softening of the caseum which is the crucial event, rather than the formation of the caseum itself.

The next stage in the evolution of the lesions is tubercle formation. The tubercles form at the peripheries of the caseum and consist of giant cells and epithelioid cells. The giant cells are derived from epithelioid cells by atypical division or fusion and may have thirty or more nuclei. Later, sclerosis may occur and this may finally be followed by calcification.

Canetti performed detailed bacillary counts on sections showing each stage of the evolution of a lesion. In the pre-exudative phase (capillary vasodilation, alveolar swelling) tubercle bacilli were numerous, and mainly seen in macrophages, often in clumps. In the exudative phase, the numbers of bacilli seen depended on the numbers of white cells present: if the alveolitis was almost purely fibrinous bacilli were rarely seen, but if the alveolitis was 'polymorphonuclear' then many bacilli were present. At the onset of caseation, the sub-type of exudative lesion again determined the amount of bacilli, with caseation arising from fibrino-macrophagic or polymorphonuclear alveolitis showing a huge increase in the number of bacilli, but caseation from a fibrinous lesion having very few bacilli.

Generalised homogeneous caseation followed by necrosis revealed bacilli only very rarely. Once necrosis had set in the bacilli appeared to die and they progressively disappeared from solid caseum. Similarly, at the periphery of the caseum, in the tubercle, few bacilli were seen even in the giant cells. Sclerotic and calcified lesions also contained no visible bacilli. However, if the caseum softened, explosive growth of bacilli occurred, surpassing the numbers seen in any other lesion and approaching the density of bacilli in pure culture.

### 7.1.3 Location of *M. tuberculosis* in tissue sections

The location of *M. tuberculosis* in infected tissues remains the subject of debate. The mystery of why so few organisms are seen in the tubercles and giant cells has prompted some to consider the possibility of a non-acid fast form of *M. tuberculosis*, perhaps in a spore-like form (the organism has genes such as *sigF* which is homologous to sporulation factors in *Streptomyces coelicolor* and *Bacillus subtilis*). Others argue that lack of visible acid-fast bacilli is simply down to the lack of sensitivity of histologic microscopy. Seiler *et al* provided persuasive evidence of a non-acid-fast form (Seiler *et al* 2003). By analysing tissue sections from both mice and humans infected with *M. tuberculosis* Seiler found that whereas ZN-staining bacilli were lost during persistence of infection, staining with polyclonal antibodies against mycobacterial antigens persisted, and bacterial titres remained constant. The authors suggest that this is best explained by mycobacterial cell-wall alterations - either loss or re-organization of cell-wall components, which might lead to the loss of the ability of the cell to retain carbol-fuchsin, or, might prevent access of the stain to the cell.

Other groups have sought to improve knowledge of the site of the bacilli by using methods to detect nucleic acids in tissues in addition to acid-fast staining. One such study (Hernandez-Pando *et al*, 2000) found evidence of *M. tuberculosis* DNA by *in-situ* and conventional PCR in 15 of 47 lung specimens taken from Mexican and Ethiopian patients who died of causes other than TB. There was no histologic evidence of the presence of *M. tuberculosis*, either as histologic lesions or as visible acid-fast bacilli. These individuals were considered to harbour latent tuberculous infection. Furthermore, the DNA was found to be present not only in macrophages, but also in type II pneumocytes, endothelial cells and fibroblasts. Thus the absence of visible acid-fast bacilli on microscopy certainly does not seem to preclude their presence in the tissues in

some form. Fenhalls *et al* (Fenhalls 2002a and 2002b) localised mycobacterial DNA and mRNA in human tuberculous granulomas. They considered viable bacilli to be those with detectable *rpoB* mRNA transcripts, with the presence of DNA alone indicating non-viable cells. In three of nine patients tuberculous granulomas were present but no acid-fast bacilli could be seen. In the other six cases, those bacilli detected by ZN staining were only found in the necrotic regions of granulomas, whilst DNA was found by *in-situ* hybridization to be present also in giant cells and macrophages on the edges of granulomas. *rpoB* mRNA was also detected in the cytoplasm of giant cells and macrophages on the outer edges of certain necrotic granulomas and a few granulomas without caseous necrosis. It was not detected within the necrotic regions of granulomas. The authors pointed out that it was possible that viable bacteria were present in other areas of the granulomas but that *rpoB* mRNA was at undetectable levels using this technology.

In addition to detection of nucleic acids, evidence of mycobacterial infection has also been sought by staining for specific mycobacterial antigens in human tissues. As already mentioned, Seiler (Seiler 2003) used a polyclonal anti-*M. bovis* BCG serum and found that whereas ZN staining gradually became negative with persistence of infection, staining with the anti-serum remained positive. Another study of human intracranial tuberculomas found that none of 10 specimens of tuberculous granuloma were ZN-positive, but mycobacterial antigens were found in the cytoplasm of giant cells and macrophages in all the cases (Radhakrishnan 1991). Hence the presence of residual antigen deposits in the tissue might provide evidence of mycobacterial infection when bacilli are not visible by ZN staining.

## 7.2 Method

### 7.2.1 *Determining the correct dilution of secondary antibodies*

For a full description of the method and antibodies used see Methods 2.7. Purified polyclonal anti-Rpf antibodies raised in sheep were obtained from Mike Young (Mukamolova et al 2002b). The secondary antibodies used were rabbit anti-sheep antibodies (Jackson ImmunoResearch). The first step was to determine the correct working dilution of rabbit anti-sheep secondary antibodies. This was done by binding various dilutions of the antibodies to sheep anti-factor 8 (Serotec) bound to sections of human placenta. The secondary antibody was tested at dilutions of 1:1000 and 1:5000, against dilutions of the sheep anti-factor 8 (primary) antibody of 1:100, 1:500, 1:1000 and negative control (no primary antibody). The outcome is shown in Table 7.1. The negative control did not stain; staining of all sections using the 1:5000 dilution of secondary antibody were too faint; the best results were obtained using 1:1000 dilution of secondary antibodies against 1:500 and 1:1000 of primary sheep anti-factor 8 antibody.

**Table 7.1**

Dilution of primary antibody (sheep anti-factor 8)	Dilution of secondary antibody (rabbit anti-sheep)	
	1:1000	1:5000
1:100	+++	+/-
1:500	+	+/-
1:1000	+	+/-
Negative control	-	-



### *7.2.2 Determining the optimal dilution of anti-Rpf antibodies and immuno-staining of tissue sections*

Sheep anti-factor 8 antibodies were replaced with sheep anti-Rpf antibodies, and instead of placental tissue sections, sections of a lymph node from the gut, known to contain *M. tuberculosis*, were used. The method was again as described in Methods 2.7. To determine the optimal dilution of anti-Rpf antibodies, a dilution of 1:1000 rabbit anti-sheep (secondary) antibody was used against the following dilutions of anti-Rpf antibodies: 1:100, 1:500, 1:1000, 1:2000. All were satisfactory but the 1:100 dilution of anti-Rpf antibodies gave the clearest result and was the dilution employed in all the figures shown. All further experiments described here were performed on consecutive sections of the same gut lymph node, unless otherwise stated. This came from a patient with known tuberculosis of the gastro-intestinal tract, and the sections were known to contain acid-fast bacilli. Features such as lymphoid follicles, granulomas and caseous lesions were all clearly visible on each slide of this piece of tissue. As all the sections were taken consecutively from the same tissue their appearances could be easily compared and the same lesion could be studied using different stains.

### *7.2.3 Absorption experiment using recombinant RpfB*

In order to confirm that the staining in the lymph node sections could be ascribed to binding of anti-Rpf antibodies to Rpf in the tissues, rather than a non-specific binding effect, freshly prepared recombinant Rpf was added to the diluted anti-Rpf antibodies prior to using them. This should neutralise the anti-Rpf antibodies and lead to a negative result. The Rpf and the anti-Rpf antibody mixture was allowed to stand at room temperature for 75 minutes before being added to the sections and immuno-staining carried out exactly as before.

#### 7.2.4 Staining of giant cells in sections not infected with *M. tuberculosis*

To rule out the possibility that giant cells (rather than Rpf) cross-reacted with the anti-Rpf antibody, sections of liver were provided by Dr A. Quaglia, Department of Histopathology, Royal Free Hospital. The sections were taken from a patient with a liver tumour, the blood supply to which had been interrupted by the procedure of embolisation (i.e. intravascular obstruction of the blood vessels, in this case deliberately, to kill the affected tissue). Later, the embolised area of liver had been resected. The embolisation had led to an inflammatory reaction in the liver, in which macrophages had fused to form giant cells, histologically very similar and comparable to those seen in tuberculous lesions. There was no clinical suspicion or evidence that the patient was infected with *M. tuberculosis*. The sections of liver were stained with anti-Rpf antibodies in exactly the same way as the sections of gut lymph node.

#### 7.2.5 Ziehl-Neelsen (ZN) staining of sections

Cold ZN staining of sections was performed in order to demonstrate whether there was any correlation between the sites in the sections which contained acid-fast bacilli and those which stained positive for Rpf. The immuno-staining was performed first, followed by ZN staining. Detail of the cold ZN staining method used is described in Methods 2.7.3.

#### 7.2.6 Staining with anti-LAM antibodies

To compare the distribution of Rpf *in-situ* with that of another mycobacterial antigen, anti-lipoarabinomannan (LAM) antibodies were obtained from Vladimir Koulchin's laboratory at Chemogen, Portland USA. LAM is a mycobacterial envelope glycolipid (Besra *et al* 1994) that has been implicated in the virulence and pathogenesis of *M. tuberculosis*. The antibodies were raised in rabbits and conjugated to horseradish

peroxidase. They were used at a 1:50 dilution, and with peroxidase-conjugated streptavidin as used in the anti-Rpf experiments. The staining method used was identical to that already described (see Methods 2.7).

### 7.3 Results

#### 7.3.1 *Staining of M. tuberculosis-infected human tissue with anti-Rpf antibodies*

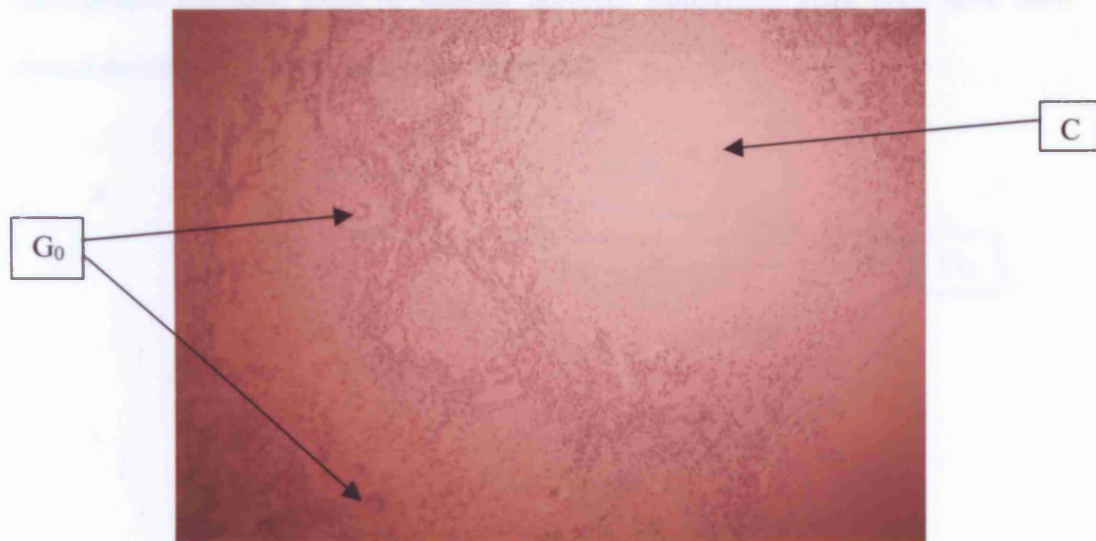
The lymph node section contained necrotic lesions with caseous centres surrounded by a zone of giant cells and macrophages. Lymphoid follicles were also visible. There was no staining of the negative control section to which anti-Rpf antibody was not applied. Figure 7.1 shows the overall appearance of the section without immuno-staining (negative control) at x400 magnification. Areas of caseating necrosis are seen surrounded by zones of giant cells and macrophages. Figure 7.2a is a x1000 magnification of epithelioid giant cells in the negative control ( $G_0$ ). In sections to which anti-Rpf antibody was applied, marked immuno-staining was seen, which appeared clearly in the giant cells and macrophages in the zones of tubercle surrounding the caseous lesions. Figures 7.2b and c show similar giant cells to those in negative control Figure 7.2a, but with anti-Rpf antibodies applied. Dark brown immunoperoxidase staining of these cells is clearly visible ( $G_1$ ). Figures 7.3a and b are x1000 magnifications of a lymphoid follicle: Figure 7.3a is the negative control and 7.3b, to which anti-Rpf antibodies were applied, again shows immunoperoxidase staining within the follicle. The images are all consecutive sections of the same infected gut lymph node.

#### 7.3.2 *Absorption experiment with recombinant Rpf*

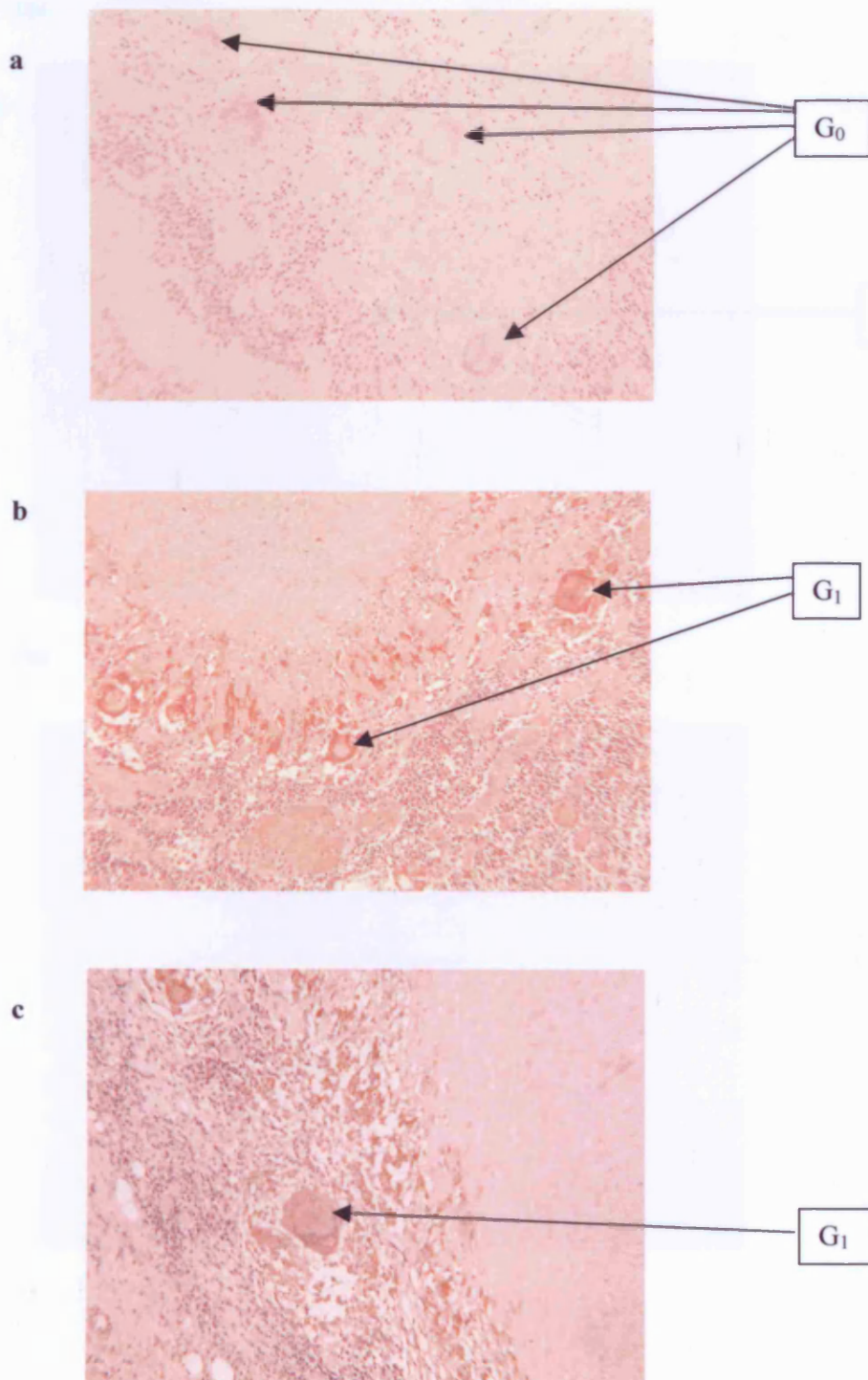
Figures 7.4a and b show the results of the absorption experiment, in which recombinant Rpf had been added to the anti-Rpf antibodies prior to staining. No staining of the

section was seen when Rpf had been added, confirming that the addition of Rpf had neutralised the anti-Rpf antibodies. Therefore the staining seen with anti-Rpf antibodies was specific to those antibodies.

**Figure 7.1** Section of *Mycobacterium tuberculosis*-infected human gut lymph node, unstained negative control, demonstrating areas of caseating necrosis (C) and giant/epithelioid cells ( $G_0$ ). x400 magnification.



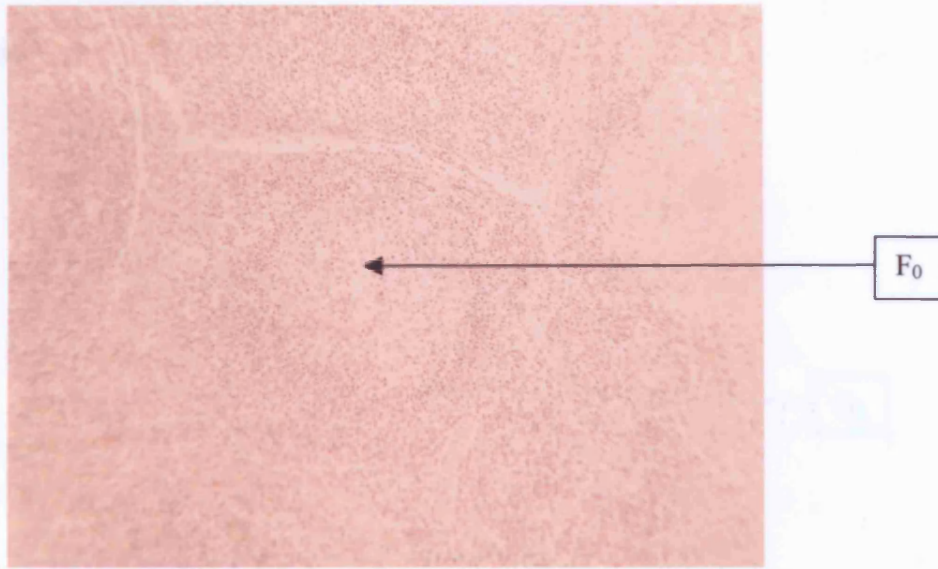
**Figure 7.2** Section of human gut lymph node (the same as in Fig 7.1) but at x1000 magnification (a) negative control showing unstained epithelioid giant cells ( $G_0$ ). (b) and (c) immuno-stained with anti-Rpf antibodies. Zones of epithelioid giant cells and macrophages surround areas of caseous necrosis. Epithelioid giant cells have been stained dark brown with immunoperoxidase ( $G_1$ ).



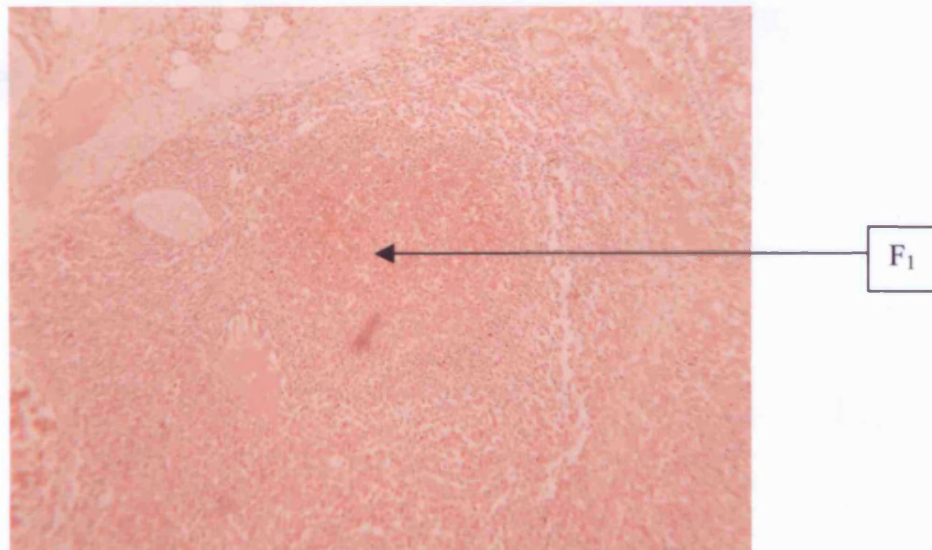
**Figure 7.3** Lymphoid follicle in section of *M. tuberculosis*-infected human gut lymph node. (a) F<sub>0</sub>, negative control; (b) F<sub>1</sub>, immuno-stained with anti-Rpf antibodies. The lymphoid follicle displays brown immunoperoxidase stain after anti-Rpf has been applied.

x1000 magnification.

(a)



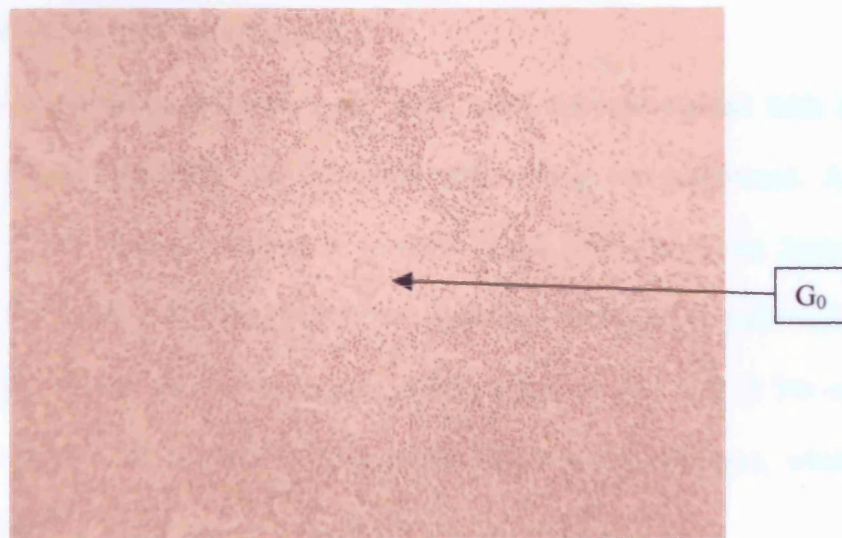
(b)



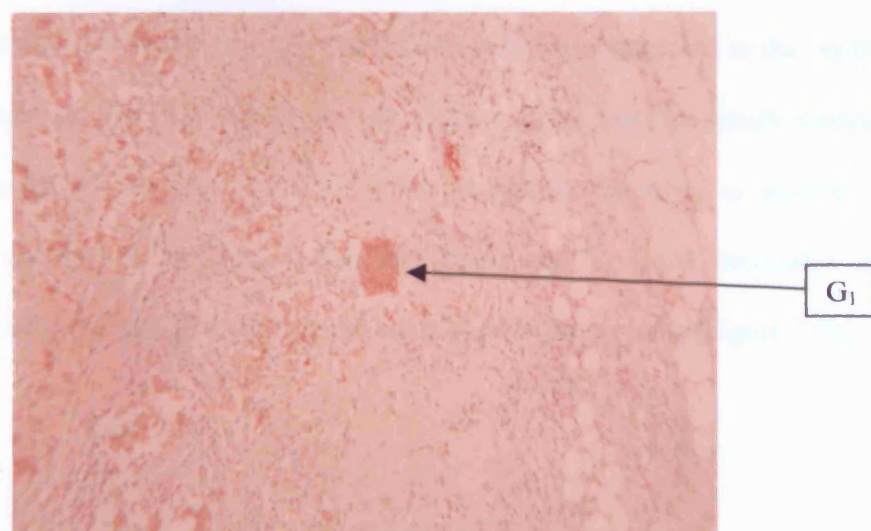


**Figure 7.4** Section of *M. tuberculosis* – infected human gut lymph node at x1000 magnification. Blocking experiment in which fresh recombinant Rpf was added to the anti-Rpf antibodies prior to immuno-staining. (a) immuno-staining after adding Rpf: there is no immuno-staining of giant cells ( $G_0$ ) and macrophages, as the anti-Rpf has been neutralised by the Rpf. (b) control in which no Rpf was added: a clearly immuno-stained giant cell is shown ( $G_1$ ).

(a)



(b)





### *7.3.3 Staining of giant cells in sections not infected with M. tuberculosis*

The appearance of the embolised liver with or without anti-Rpf antibodies was the same. The giant cells in the liver sections did not stain in the presence of anti-Rpf antibodies (figures 7.5a and b). This ruled out the possibility that giant cells cross-reacted with the anti-Rpf antibody. A positive control made using the gut lymph node section did show immuno-staining.

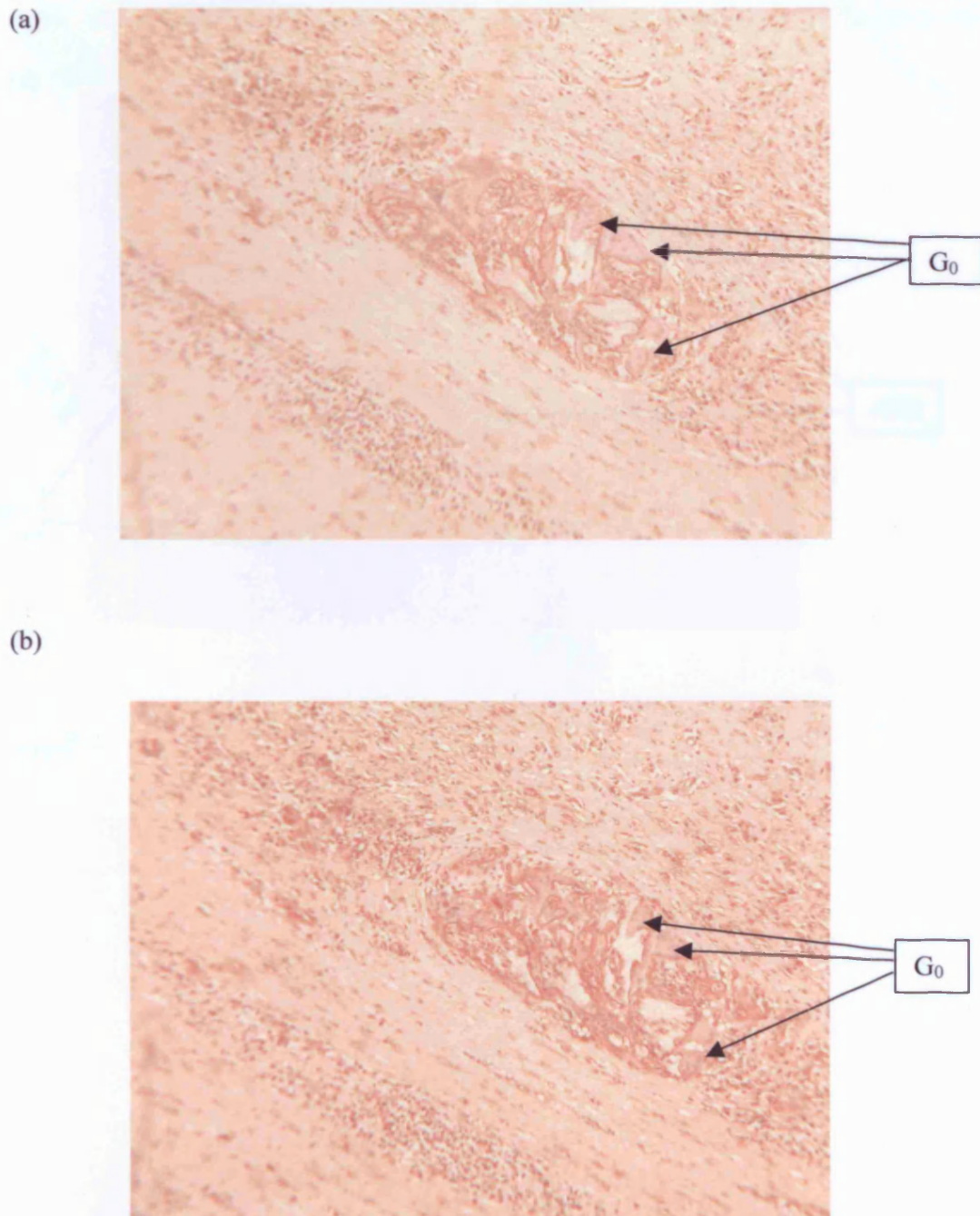
### *7.3.4 Ziehl-Neelson staining of sections*

The sections of human gut lymph node were again immuno-stained with a 1:100 dilution of anti-Rpf antibody and then cold ZN-staining was performed. Acid-fast bacilli were visible only in the centres of some of the necrotic lesions. Immediately surrounding the bacilli were areas of immuno-staining with anti-Rpf antibodies, which was not present in the rest of the necrotic lesion (Figures 7.6a and b). No acid-fast bacilli were seen in the areas containing giant cells and macrophages, which were positive for Rpf.

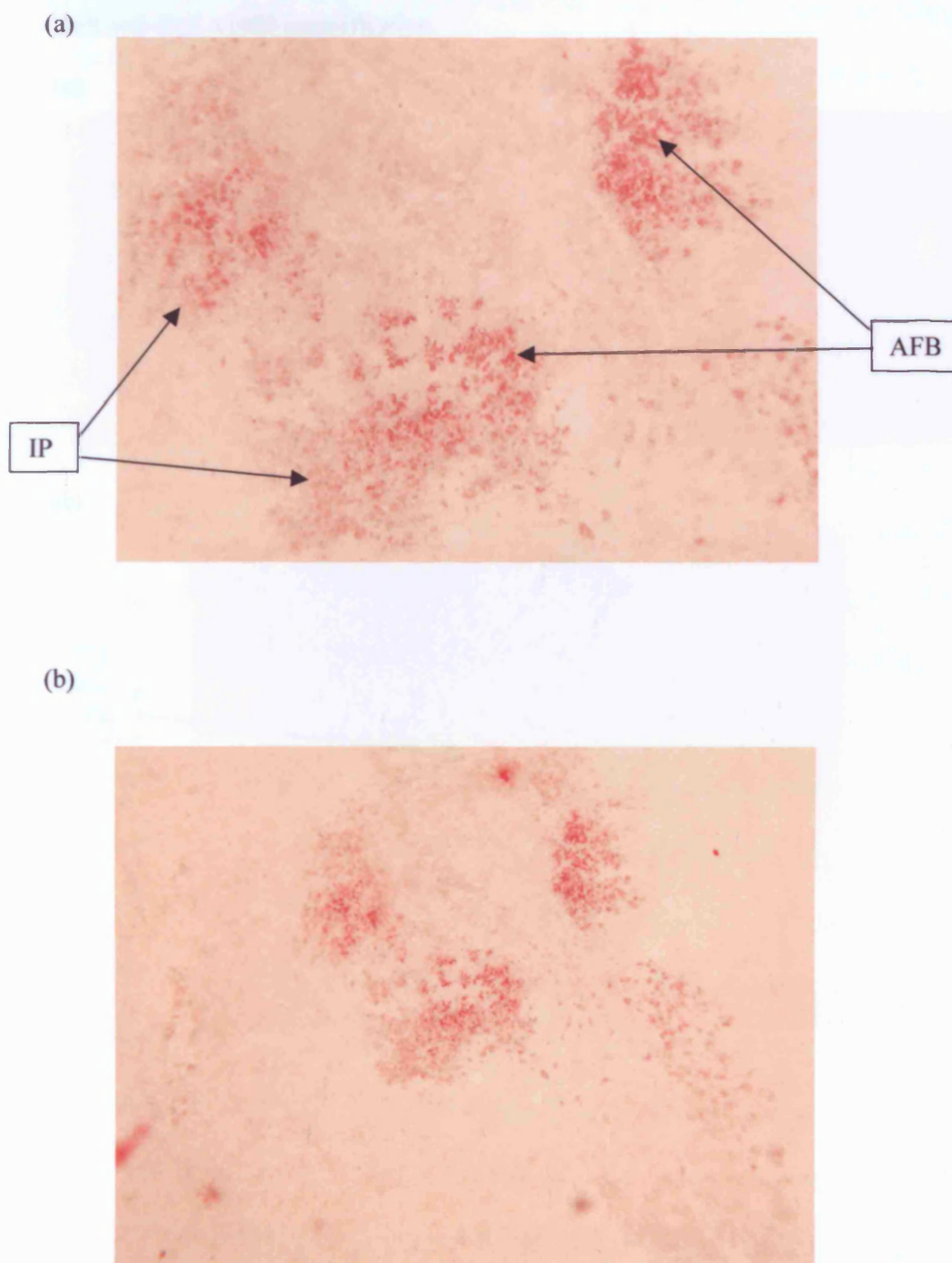
### *7.3.5 Staining with anti-LAM antibodies*

Staining with anti-LAM antibodies revealed that LAM was detectable at the centres of some of the necrotic lesions in the immediate vicinity of the acid-fast bacilli themselves (Figure 7.7 a and b), in a similar distribution to Rpf as described in section 7.3.4. However in contrast to the pattern with Rpf there was no LAM detectable in the macrophages or giant cells surrounding the areas of caseous necrosis (Figure 7.7c).

**Figure 7.5** Giant cells ( $G_0$ ) in a section of embolised human liver, uninfected with *M. tuberculosis*, magnification x1000. (a) negative control without addition of anti-Rpf antibodies (b) the same section immuno-stained with anti-Rpf antibodies. There is no immuno-staining of giant cells with anti-Rpf antibodies, indicating that immuno-staining is not due to a non-specific cross-reaction with giant cells.

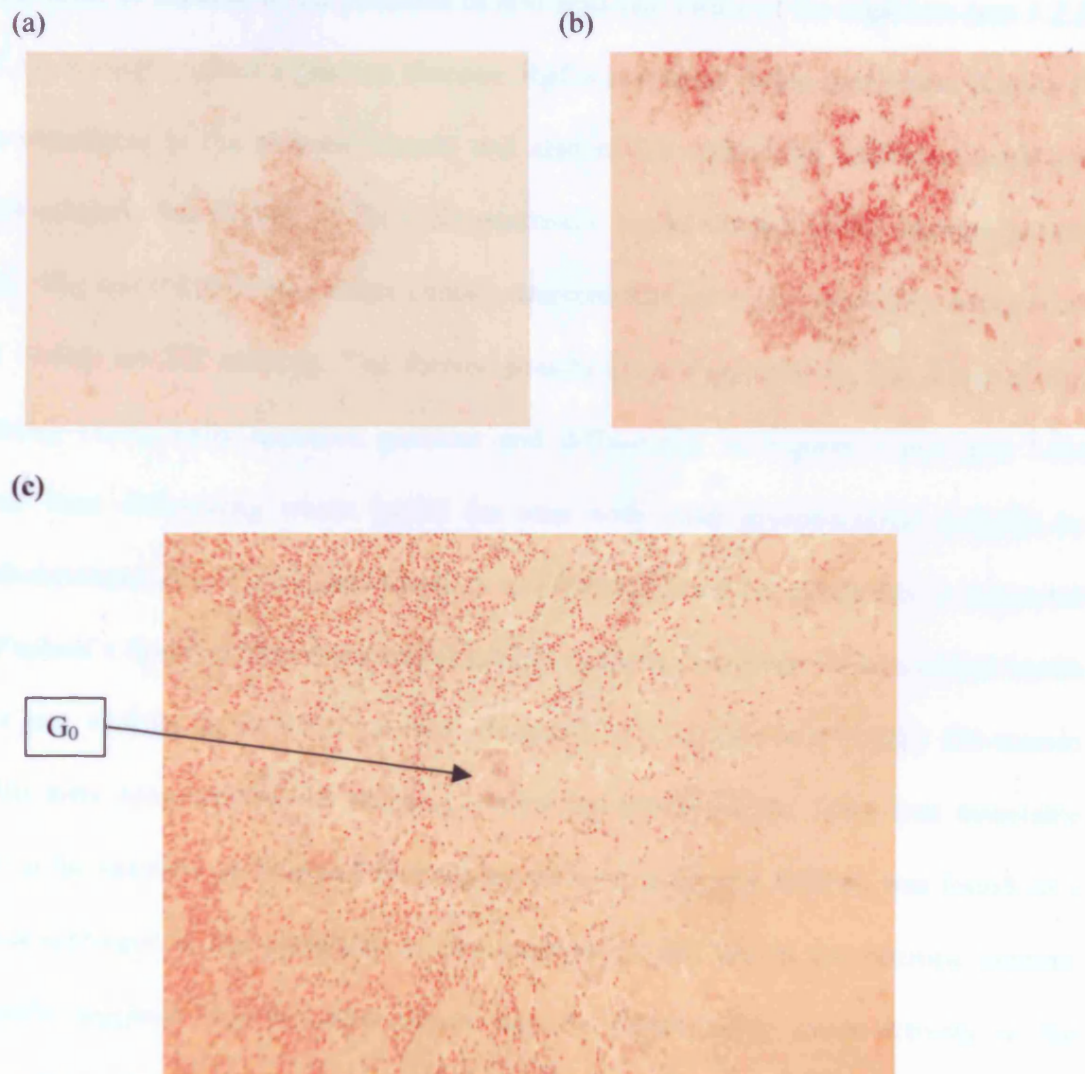


**Figure 7.6** Area of caseous necrosis in a section of *M. tuberculosis*-infected human gut lymph node, after Ziehl-Neelsen staining and immuno-staining with anti-Rpf, magnification x1000. (a) acid-fast bacilli (AFB) are seen surrounded by areas of brown immunoperoxidase staining (IP). (b) the same section, showing the absence of immunoperoxidase staining in the necrotic areas distant from acid-fast bacilli.





**Figure 7.7** (a) Same area of section as seen ZN-stained in Figure 7.6 but immuno-stained with anti-LAM antibodies. Staining is seen around the areas where acid-fast bacilli are located, in a similar way to staining with anti-Rpf. (b) same section with ZN and anti-Rpf immuno-staining for comparison. (c) Same section, at edge of area of caseous necrosis, including giant cells (G), after staining with anti-LAM. There is no immuno-staining, in contrast to the marked staining of giant cells and macrophages seen with anti-Rpf. x1000 magnification.

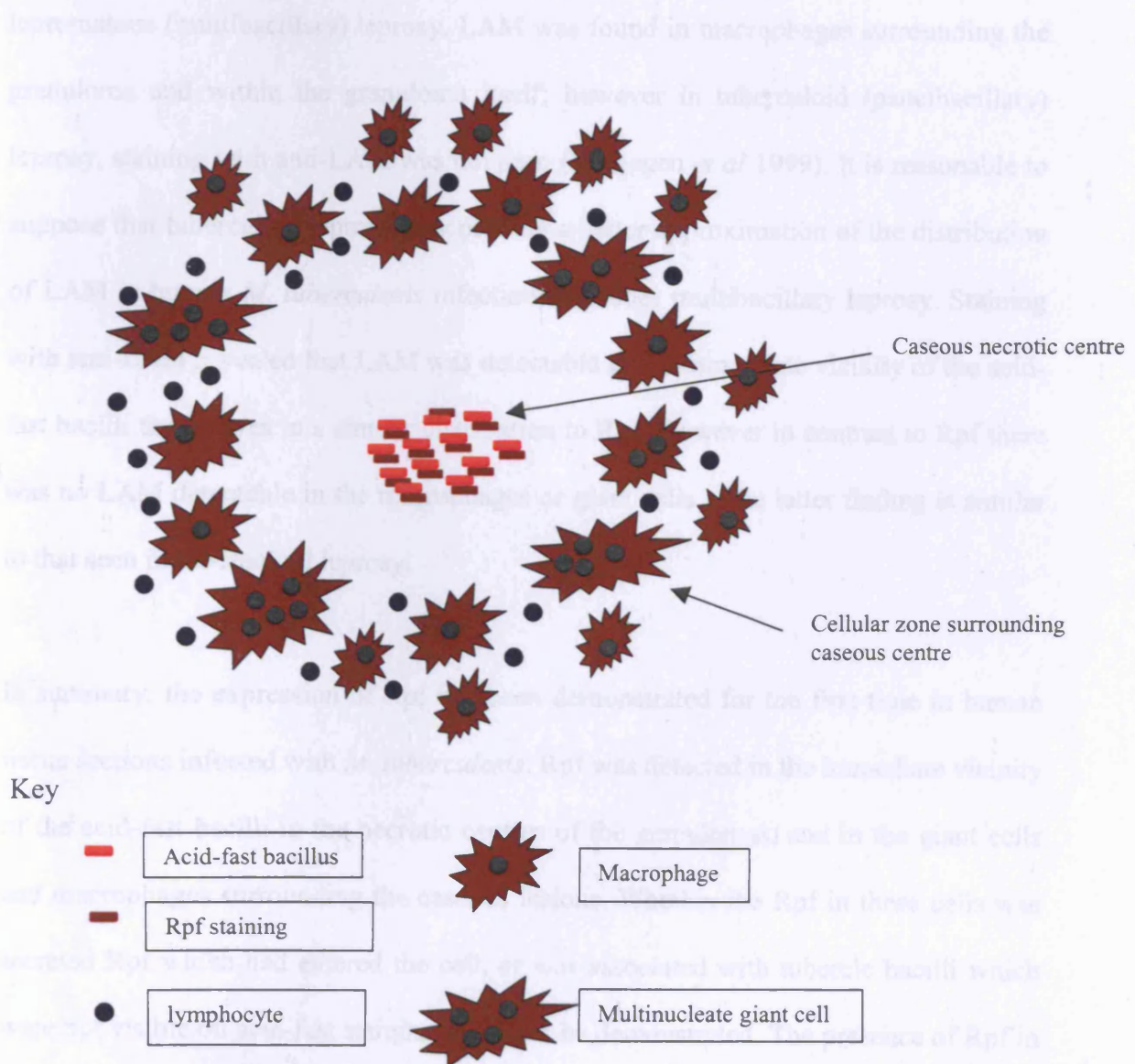


## 7.4 Discussion

This is the first time anti-Rpf antibodies have been applied to human tissue infected with *M. tuberculosis*. Rpf was detected in human tissue for the first time. Acid-fast bacilli were present on ZN-staining only in the necrotic centres of the tuberculous lesions. The apparent absence of acid-fast bacilli from the macrophages and giant cells might be because of the insensitivity of acid-fast staining at detecting low numbers of organisms, or because of the presence of non-acid-fast forms of the organism (see 1.2.2; 7.1.3), or might reflect a genuine absence. Rpf was present in the immediate vicinity of the organisms in the necrotic centres and also in the epithelioid cells/giant cells and macrophages. See Figure 7.8 for a diagrammatic representation of the staining pattern seen. The stained cells may either contain secreted Rpf, or entire organisms which were not visible on ZN staining. The former possibility is supported by the fact that Rpf staining within cells appeared granular and diffuse (G<sub>1</sub> in Figures 7.2b/c and 7.4b) rather than delineating whole bacilli (as seen with other mycobacterial antigens by Radhakrishnan *et al* (1991) and Orrell *et al* (1991)). The latter possibility is supported by Fenhall's findings of mycobacterial mRNA in these cells even when acid-fast bacilli were not visible. In Fenhall's studies (Fenhalls *et al* 2002a and 2002b) ZN-stained bacilli were seen only in the necrotic centres but mycobacterial DNA was detectable both in the necrotic centres and in the giant cells/macrophages. mRNA was found only in macrophages at the periphery of the granulomas and not in the necrotic centres. Fenhalls suggests that *M. tuberculosis* may be metabolising more actively in the macrophages than in other areas of the granulomas, and that mRNA may also be present at undetectable levels in other areas such as the necrotic centres. This scenario would fit with the results presented here, where most Rpf was detected in the giant cells/macrophages and a smaller amount associated with the ZN-stained bacilli in the necrotic area.

**Figure 7.8**

Diagrammatic representation of a granuloma in human tissue infected with *M. tuberculosis*, showing the distribution of immuno-peroxidase staining after immuno-staining with anti-Rpf antibody. Areas shown in brown are immuno-stained. Rpf is present in macrophages and multinucleate giant cells, and also in the immediate vicinity of acid-fast bacilli in the caseous necrotic centre.



LAM is a mycobacterial envelope glycolipid (Besra *et al* 1994) that has been implicated in the virulence and pathogenesis of *M. tuberculosis* and *M. leprae* infection and has been associated with a variety of deleterious effects on immune system function. The *in-situ* expression of LAM in human tuberculous disease does not appear to have been studied although several studies have looked at its expression in human leprosy (Verhagen *et al* 1999, van den Bos *et al* 1999, Lockwood *et al* 2002). In skin lesions of lepromatous (multibacillary) leprosy, LAM was found in macrophages surrounding the granuloma and within the granuloma itself; however in tuberculoid (paucibacillary) leprosy, staining with anti-LAM was not seen (Verhagen *et al* 1999). It is reasonable to suppose that tuberculoid leprosy may provide a better approximation of the distribution of LAM in human *M. tuberculosis* infection than does multibacillary leprosy. Staining with anti-LAM revealed that LAM was detectable in the immediate vicinity of the acid-fast bacilli themselves in a similar distribution to Rpf. However in contrast to Rpf there was no LAM detectable in the macrophages or giant cells. The latter finding is similar to that seen in tuberculoid leprosy.

In summary, the expression of Rpf has been demonstrated for the first time in human tissue sections infected with *M. tuberculosis*. Rpf was detected in the immediate vicinity of the acid-fast bacilli in the necrotic centres of the granulomas, and in the giant cells and macrophages surrounding the caseous lesions. Whether the Rpf in these cells was secreted Rpf which had entered the cell, or was associated with tubercle bacilli which were not visible on acid-fast staining could not be demonstrated. The presence of Rpf in human tuberculous infection lends support to the idea that Rpf has a role in controlling dormancy in human tuberculosis.

## **CHAPTER VIII**

### **DISCUSSION**



## 8.1 Discussion

If medicine is ever to be successful in controlling tuberculosis, then the problem of treating persisters and latent infection must be overcome. Several new types of interventions could help target latent tuberculosis. These include improved diagnostic testing for latent infection, to replace the tuberculin skin test which has poor sensitivity and specificity; an effective post-exposure vaccine that would prevent subsequent reactivation – already achieved with a DNA vaccine in mice (Lowrie *et al* 1999); and the development of drugs which can effectively and rapidly kill persisters and tubercle bacilli in the dormant state. To achieve all these, a better understanding of the dormancy response is necessary. Such knowledge could help transform the current approaches to tuberculosis treatment.

In order to identify specific latency drug targets, several lines of research are underway. For example there is a major effort to develop rationally designed inhibitors of the isocitrate lyase enzyme, and studies of the action of sigma factor could lead to rational inhibitors of a latency-specific transcription factor. If the biochemical triggers for reactivation from the dormant state were understood, perhaps this would be prevented in latent tuberculosis; on the other hand, if growth deceleration and latency could be blocked then bacilli in the host could be more readily killed with existing drugs. Therefore the identification of molecules which caused the reactivation of dormant cells of *M. tuberculosis* was greeted with great interest. As well as Rpfs, phospholipids and other small peptides have been implicated as having a role in *M. tuberculosis* reactivation (Zhang *et al* 2001). This group noted that spent culture supernatant from stationary phase cultures of *M. tuberculosis* could resuscitate dormant cells of the bacillus. They found evidence that a phospholipid and a small 8 kDa peptide, Rv1174c, were responsible for this activity, and pointed out that exogenous phospholipids could serve as building materials for the cell membrane during their

replication or repair. Alternatively they might have a role as signalling molecules for growth.

The concept of signalling between bacteria is not a new one. Communities of bacteria convey their presence to each other by releasing and detecting chemical signalling molecules called autoinducers. This is known as quorum sensing. Many bacterial behaviours are regulated by quorum sensing, such as symbiosis, virulence, antibiotic production and biofilm formation. The first- and best-known example is the case of the ocean-dwelling Gram-negative organism *Vibrio fischeri*. This organism constantly secretes an autoinducer which at low concentrations has no effect, but which at high concentrations activates the *lux* operon, resulting in the production of luciferase and the emission of light by the organisms. This therefore only occurs when the bacteria are present in large numbers, so that energy is not wasted on light production when there are too few present to have an effect. The emission of light gives the bacteria an advantage as it results in a symbiotic relationship with a squid, which carries the organisms in a light organ. Two regulatory components are required for the process. The LuxI protein is responsible for production of the homoserine lactone (HSL) autoinducer, and the LuxR protein is responsible for binding the HSL autoinducer and activating transcription of the luciferase structural operon at high density. HSL/transcriptional activator signalling circuits have since been found to be the standard mechanism by which many Gram-negative bacteria communicate (Schauder & Bassler 2001). Other organisms exhibiting inter-cellular signalling using this method include *Pseudomonas aeruginosa* (to regulate expression of virulence and biofilm formation). Quorum sensing in Gram-positive bacteria, however, uses modified oligopeptides that are secreted into the medium and accumulate at high cell density. The detectors for these oligopeptide signals are two-component adaptive response

proteins, and the mechanism of signal transduction is via a conserved phosphorylation/dephosphorylation mechanism, as described in section 6.7. This mechanism of action was thought the most likely for Rpf until recently.

Whilst inter-bacterial communication is thus well-described, the idea of a bacterial ‘growth factor’ is novel. Since Rpfs are present in sporulating organisms such as streptomycetes it was suggested that Rpfs might be the trigger for the organism to come out of the sporulated state. Sporulation is itself a protective response to environmental conditions: when the environmental conditions are again favourable for growth, some organisms in the population might start secreting Rpfs, which might reactivate others in the population. This scenario could be extended to apply to the dormant state in genetically related though non-spore-forming organisms. However the theory leaves many questions unanswered. How are the first organisms to start secreting Rpf reactivated? How do they sense that conditions have become favourable for growth? Presumably at least one organism must start reactivating and secreting Rpf without such a signal. Since Rpfs seem to have cross-species reactivity then in an environment such as soil, when one species begins to secrete Rpf, organisms of other species would also be reactivated – a situation not necessarily beneficial to the first organism as there would then be more competition for nutrients.

Recent bioinformatic work has indicated that the high G+C Gram-positive organisms may not have a monopoly on Rpf-like proteins (Ravagnani, Finan & Young 2005). Although the Rpf domain itself is not found in low G+C Gram-positives, an alternative domain dubbed the Sps (stationary phase survival) domain has been identified, which has very similar accessory domain structures and genomic contexts, for example in *Clostridium* sp., *Bacillus* sp., *Listeria* sp. and *Enterococcus faecalis* (see Introduction 1.4.2). It may even be that

proteins with a similar function to Rpf are near-universal amongst bacteria, and that they may have clinical relevance in many other diseases as well as tuberculosis: if so, for example, inhibiting the action of the Rpf-equivalent proteins might prevent proliferation of bacteria in the human host.

The discovery of resuscitation-promoting factors in *M. tuberculosis* offers a lead in unravelling the mechanisms of dormancy and reactivation in this organism, as well as a potential target for drug therapy. However, since 1998, when Rpfs were first identified within the *M. tuberculosis* genome, progress in characterising them and finding their mode of action has been frustratingly slow. This has mainly been a result of their instability and insolubility in recombinant form, making work technically difficult and time-consuming even for those most experienced in the field. Commercial companies, interested in the potential of Rpfs at first, have for these reasons ceased to prioritise further work on Rpfs (Ken Duncan, personal communication). Recent progress in identifying a possible mode of action has relied almost exclusively on a bioinformatic approach.

Much of the work described in this thesis was aimed at improving the expression and solubility of mycobacterial (*M. tuberculosis* and *M. smegmatis*) Rpfs, using various approaches. The best success was achieved in the case of the small Rpf of *M. smegmatis* (Rpf<sub>smeg</sub>). By expressing Rpf<sub>smeg</sub> as a fusion protein with NusA, a much greater quantity of soluble protein could be extracted, without the use of a denaturant. Previous attempts to express this protein had resulted in quantities too small to visualise on immunoblotting. Quantities sufficient for testing of biological activity were now produced for the first time. The use of a fusion protein with Rpf<sub>smeg</sub> was more successful than the strategy of using

lower induction temperatures and lower concentrations of inducer with RpfB and RpfE, although this too yielded some soluble protein.

The next step was to devise a convenient test for the biological activity of the Rpfs, so that batches of recombinant protein could be shown to be suitable for use in further experiments. In order to do this, suitable bacterial cultures containing cells in a dormant state were required. Standardised cultures of *M. bovis* BCG were developed which were reliable and gave useful results. The activity of Rpf<sub>smeg</sub>, hitherto untested, was assessed using these cultures. Results confirmed it was active, in the fM-pM range. The tiny Rpf<sub>smeg</sub> consists only of the core 'Rpf domain' common to all Rpfs, and these experiments were the first time that this domain alone, in a naturally-occurring (as opposed to artificially truncated form) had been tested for activity. The results confirmed that this domain is the active region in the Rpf proteins.

Despite being reliable, cells from this model of dormancy in *M. bovis* BCG took between 3 and 6 months to prepare, and from 1 to 3 weeks to give results in an activity assay. Therefore attention turned to *M. luteus*, a much more rapidly growing organism. Straightforward optical density measurements were not accurate enough because of the tendency of the cultures to clump, and because they were not sensitive enough at the very low inoculum density required to demonstrate an effect of Rpf. By measuring the ATP concentration of cultures these problems were overcome and rapid (4 hour culture) and overnight (16 hour culture) tests of Rpf activity were developed. The latter allowed accurate quantitative comparisons of the activity of different preparations of recombinant Rpf, whilst the former was more suitable as a rapid qualitative test of whether a batch had activity or not. The range of activity of RpfB against *M. luteus* had not been measured

previously. The most active concentration of RpfB was in the nanomolar range. Most other Rpfs studied have had optimal activity in picomolar concentration (Mukamolova 2002b). Since RpfB, unlike the other Rpfs, is thought to be membrane bound rather than secreted, its optimal concentration for activity might also differ; or, it could simply be that the different method used to purify the protein in these experiments (commercial rather than in-house columns, with different buffers) might in some way have affected the protein folding process or its stability (Mukamolova's work found that RpfB in picomolar concentrations was active against *M. bovis* BCG).

The mode of action of Rpfs is not yet fully known although bioinformatic analysis over the past 12 months has begun to shed some light on this question. Two possible mechanisms have been considered. The most obvious is that the Rpfs act in an autocrine manner and bind to a very high affinity cell surface receptor. This could be part of a two-component sensory recognition apparatus, the model that would be intuitively expected when drawing comparisons with other examples of quorum sensing in Gram-positive bacteria: or could involve other specific protein receptors, such as occurs in eukaryotes, in which growth factors bind tyrosine kinases. An alternative explanation is that the Rpfs act as enzymes and either by cleaving the cell wall, or by cleaving the cell wall and releasing the key mediator, bacteria enter the cell cycle. The first hypothesis was tested here by means of a phage display system.

Panning of the phage display library against recombinant RpfB identified an identical 400 bp fragment in 25% of the clones which contained a 17 bp sequence matching a sequence in *M. tuberculosis* gene Rv0328. This gene is predicted to code for a transcription regulatory protein, and was successfully cloned. Unfortunately the expressed protein had

low solubility with the expression system used. Because of time constraints a different expression system could not be tried, nor binding experiments performed to see whether Rv0328 did bind Rpf. The possibility that Rv0328 does bind Rpf was therefore not excluded. Had the time been available to improve the protein's solubility, binding assays with Rpf could have been performed, for example by passing the His-tagged Rv0328 down a nickel column, followed by recombinant Rpf or supernatant of dormant *M. luteus* cultures.

Since this attempt to identify a protein receptor was made, evidence has been mounting that Rpf may in fact act as an enzyme, facilitating cleavage of the cell-wall peptidoglycan, and not by binding to a specific protein receptor at all. This was first suspected when the similarity of one of the folds of Rpf to the active site of lysozyme was noted (Cohen-Gonsaud *et al* 2004). See Introduction 1.4.2 for a review of this work. Hence, although the overall sequence conservation between Rpf and c-type lysozymes is low, it was suggested that the function of Rpfs may be to cleave oligosaccharide in the cell wall. Mike Young's group's recent publication, in which the Sps domain was also identified, supports this view (Ravagnani, Finan & Young 2005). The group suggests that the walls of dormant organisms may be considered similar to a 'cocoon', requiring the muralytic enzymes to cleave certain residues so that growth and cell wall expansion can resume. Inter-cellular signalling might still be involved, mediated perhaps by the release of a small molecule as a result of the cleavage of peptidoglycan by Rpf.

Whatever the mode of action, the role of Rpf in human tuberculosis infection – if any – remains uncertain. To discover whether or not Rpfs are in fact expressed in infected human tissue, paraffin-embedded tissue sections taken from human cases of tuberculosis were

investigated by performing immunocytochemistry with anti-Rpf antibodies. Expression of Rpf in human tissue was successfully demonstrated for the first time. Although acid-fast bacilli were present on ZN-staining only in the necrotic centres of the tuberculous lesions, Rpf appeared to be present not only in the immediate vicinity of the organisms, but also in the epithelioid cells/giant cells and macrophages in the zone surrounding the caseous necrotic lesions. These cells may either contain secreted Rpf, or entire organisms which were not visible on ZN staining. This was consistent with the results of Fenhalls' studies of mycobacterial nucleic acids in such lesions (Fenhalls 2002a and 2002b), in which ZN-stained bacilli were seen only in the necrotic centres (as here) but mycobacterial DNA was detectable both in the necrotic centres and in the giant cells/macrophages, and mRNA was found only in macrophages at the periphery of the granulomas and not in the necrotic centres. *M. tuberculosis* may be metabolising more actively in the macrophages than in other areas of the granulomas, with most Rpf/mRNA detected in the giant cells/macrophages and smaller (Rpf) or undetectable (mRNA) amounts associated with the ZN-stained bacilli in the necrotic area. The presence of Rpf in human tuberculous infection means that the possibility of a role for the Rpfs in controlling dormancy in human disease remains very much a possibility.

## **8.2 Further Work**

Technical challenges still face researchers studying Rpfs. Difficulties with solubility and measuring activity still need to be more completely overcome. Of the strategies assessed here to improve solubility, that of expression using a NusA fusion protein was by far the most successful. Production of clones expressing all the Rpfs of *M. tuberculosis* with such a fusion protein would be worthwhile, with subsequent manipulation of induction temperatures to further enhance solubility. The reliable production of sufficient quantities



of soluble Rpf s would greatly assist further efforts to assess their effects both *in vitro* and *in vivo*.

The most successful experiments described in this thesis were those involving the immunocytochemistry of human *M. tuberculosis*- infected tissue using anti-Rpf antibodies. This approach merits further work. Different human tissues should be stained, particularly lung tissue, to demonstrate the presence and location of expressed Rpf s. Tissues from patients with clinically active disease should be compared with those from cases with latent tuberculosis, to see whether there is a difference in the pattern of Rpf production between these two groups. If specific antibodies to each individual Rpf could be raised then the expression and location of each could be analysed individually. Such specific antibodies with no cross-reactivity with other Rpf s might be very difficult to produce but an alternative approach would be to stain tissues from mice infected with Rpf knock-out mutant strains of *M. tuberculosis* (Downing 2005), for comparison.

If little is known about the function of Rpf s, less still is yet known about their effect on virulence or pathogenesis in the host. Mutants of *M. tuberculosis* lacking three of the *rpf* genes (*rpfA*, *rpfB* and *rpfC* or *rpfA*, *rpfC* and *rpfD*) were significantly attenuated in a mouse infection model (Downing 2005). Further work in this area could be very enlightening. Purified recombinant Rpf s, or Rpf analogues, could be administered to mice infected with such attenuated strains of *M. tuberculosis*, and *M. tuberculosis* H37Rv, to see whether, and how, this influenced the disease process. What would be the consequence of infection with a mutant lacking all five Rpf s, if construction of such a mutant were possible? Alternatively, what would be the effect of administering anti-Rpf antibodies, or synthetic Rpf antagonists, to mice infected with wild-type *M. tuberculosis*?

The study of dormancy in general, and Rpfs in particular, remain key areas in which a better understanding may yet transform the treatment and prognosis of the one third of the world's population infected with tuberculosis. If the full mechanism of action of Rpfs could be determined, and their role in human infection characterised, there could be a choice of approaches to take in combating human disease. Either suppressing their activity to prevent reactivation of dormant organisms, or mimicking them to prevent persisters developing and to activate latent organisms and make them treatable, might be possible. Furthermore, Rpfs have potential for use in diagnostic culture of slow-growing mycobacteria. The ability to allow very small inocula to grow, and to shorten the lag phase of a culture could both be extremely helpful as ways to make rapid culture systems both faster and more sensitive. Some preliminary work in this field has been attempted (Freeman *et al* 2002). On a broader scale, Rpfs may only be one enzyme group of many similar involved in the growth kinetics of different bacteria, as is hinted at with the identification of the Sps-domain proteins. The idea of a 'cocoon'-like state with cell wall alterations, if confirmed, might for example be relevant to organisms in biofilms, where an ability to penetrate the biofilm and cleave the cell wall enzymatically might allow more effective antibiotic treatment of these refractory infections. Many common pathogens may be found to have other similar systems, which could be harnessed as an important adjunct to antibiotic therapy, and which might assume great importance in an age of ever-increasing antibiotic resistance.

## REFERENCES

## References

- Abou-Zeid C, Garbe T, Lathigra R, Wiker HG, Harboe M, Rook GA, and Young DB. (1991) Genetic and immunological analysis of *Mycobacterium tuberculosis* fibronectin-binding proteins. *Infect Immun.* **59**(8): 2712–2718.
- American Thoracic Society. (2000) Diagnostic Standards and Classification of Tuberculosis in Adults and Children *Am. J. Respir. Crit. Care Med.* **161** (4): 1376-1395
- Barer M. (2003) Physiological and molecular aspects of growth, non-growth, culturability and viability in bacteria. p.1-35. *In* ARM Coates (ed.), Dormancy and Low-Growth States in Microbial Disease. Cambridge University Press, Cambridge.
- Bateman A & Bycroft M. (2000) The Structure of a LysM Domain from *E. coli* Membrane-bound Lytic Murein Transglycosylase D (MltD). *J Mol Biol* **299**: 1113-1119
- zu Bentrup KH & Russell DG. (2001) Mycobacterial persistence: adaptation to a changing environment. *Trends in Microbiology* **9**: 597- 605
- Besra GS, Chatterjee D. (1994) Lipids and carbohydrates of *Mycobacterium tuberculosis*, p. 285-306. *In* B.R. Bloom (ed.), Tuberculosis: pathogenesis, protection and control. ASM press, Washington DC.

Betts JC, Lukey PT, Robb LC, McAdam RA, Duncan K. (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* **43**: 717-731

Biketov S, Mukamolova GV, Potapov V, Gilenkov E, Vostroknutova G, Kell DB, Young M, Kaprelyants AS. (2000) Culturability of *Mycobacterium tuberculosis* cells isolated from murine macrophages: a bacterial growth factor promotes recovery. *FEMS Immunol Med Microbiol* **29**: 233-40

Bishai W. (2000) Rekindling old controversy on elusive lair of latent tuberculosis. *Lancet* **356**: 2113 – 2114

Canetti G. (1955) The Tubercle Bacillus in the Pulmonary Lesion of Man: Histobacteriology and its Bearing on the Therapy of Pulmonary Tuberculosis. Springer Publishing Company Inc, New York.

Chambers SP, Prior SE, Barstow DA, Minton NP. (1988) The pMTL *nic*-cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* **68**: 139-49

Chiba Y & Kurihara T. (1979) Development of pulmonary tuberculosis with special reference to the time interval after tuberculin conversion. *Bull Int Union Tuberc* **54**: 263-264.

Clark E de B. (2001) Protein refolding for industrial processes. *Current Opinion in Biotechnology* **12**: 202-207

- Cohen-Gonsaud M, Keep NH, Davies AP, Ward J, Henderson B, Labesse G. (2004) Resuscitation-promoting factors possess a lysozyme-like domain. *Trends Biochem Sci* **29**: 7-10
- Cohen-Gonsaud M, Barthe P, Bagneris C, Henderson B, Ward J, Roumestand C, Keep NH. (2005) The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. *Nature Structural and Molecular Biology* **12**: 270-273
- Cunningham AF, Spreadbury CL (1998) Mycobacterial stationary phase induced by low oxygen tension: cell wall thickening and localization of the 16-kilodalton  $\alpha$ -crystallin homolog. *J Bacteriol* **180**: 801-808
- Davis GD, Elisee C, Newham DM, Harrison RG. (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol Bioeng* **65**:382-8.
- Dick T, Lee BH, Murugasu-Oei B. (1998) Oxygen Depletion-Induced Dormancy in *Mycobacterium smegmatis*. *FEMS Microbiol Lett* **163**: 159-164
- Downing KJ, Betts JC, Young DI, McAdam RA, Kelly F, Young M, Mizrahi V. (2004) Global expression profiling of strains harbouring null mutations reveals that the five rpf-like genes of *Mycobacterium tuberculosis* show functional redundancy. *Tuberculosis* **84**: 167-179
- Downing KJ, Mischenko VV, Shleeve MO, Young DI, Young M, Kaprelyants AS, Apt AS, Mizrahi V. (2005) Mutants of *Mycobacterium tuberculosis* lacking three of

the five *rpf*-like genes are defective for growth in vivo and for resuscitation in vitro.

*Infect Immun* **73**: 3038-3043

East African/British Medical Research Council (1973) Controlled clinical trial of four short-course (6-month) regimens of chemotherapy for treatment of pulmonary tuberculosis. *Lancet* **1**:1331–1339

Fenhalls G, Stevens-Muller L, Warren R, Carroll N, Bezuidenhout J, van Helden P, Bardin P (2002a) Localisation of mycobacterial DNA and mRNA in human tuberculous granulomas. *Journal of Microbiological Methods* **51**: 197-208

Fenhalls G, Stevens L, Moses L, Bezuidenhout J, Betts JC, van Helden P, Lukey PT, Duncan K (2002b) *In situ* Detection of *Mycobacterium tuberculosis* Transcripts in Human Lung Granulomas Reveals Differential Gene Expression in Necrotic Lesions. *Infect Immun* **70**: 6330-6338

Fox W, Ellard GA, Mitchison DA. (1999) Studies on the treatment of tuberculosis undertaken by the British Medical Research Council tuberculosis units, 1946 – 1986, with relevant subsequent publications. *Int J Tuberc Lung Dis* **3** (10 suppl 2): S231-79

Freeman R, Dunn J, Magee J, Barrett A. (2002) The enhancement of isolation of mycobacteria from a rapid liquid culture system by broth culture supernate of *Micrococcus luteus*. *J Med Micro* **51**: 92-93

Harrison RG. (2000) Expression of soluble heterologous proteins via fusion with NusA protein. *inNovations* **11**: 4-7

Hartmann M, Barsch A, Niehaus K, Puhler A, Tauch A, Kalinowski J. (2004) The glycosylated cell surface protein Rpf2, containing a resuscitation-promoting factor motif, is involved in intercellular communication of *Corynebacterium glutamicum*. *Arch Microbiol* **182**: 299-312

Heifets L, Lindholm-Levy. (1992) Pyrazinamide sterilizing activity in vitro against semidormant *Mycobacterium tuberculosis* bacterial populations. *Am Rev Respir Dis* **145**: 1223-5

Heilmann C, Herrmann M, Kehrel BE, Peters G. (2002) Platelet-Binding Domains in 2 Fibrinogen-Binding Proteins of *Staphylococcus aureus* Identified by Phage Display. *J Inf Dis* **186**: 32-9

Hernandez-Pando R, Jeyanathan M, Mengistu G, Aguilar D, Orozco H, Harboe M, Rook GA, Bjune G. (2000) Persistence of DNA from *Mycobacterium tuberculosis* in superficially normal lung tissue during latent infection. *Lancet* **356**: 2133-2138

Hu Y, Coates ARM. (2003) Tuberculosis, p.181-207. In ARM Coates (ed.), Dormancy and Low-Growth States in Microbial Disease. Cambridge University Press, Cambridge.

Hu Y, Mangan JA, Dhillon J, Sole KM, Mitchison DA, Butcher PD, Coates ARM. (2000) Detection of mRNA Transcripts and Active Transcription in Persistent *Mycobacterium tuberculosis* induced by exposure to Rifampicin or Pyrazinamide *J Bact* **182**: 6358-6365



Heubschmann P. (1929) Pathologische Anatomie der Tuberkulose. J. Springer, Berlin.

Jacobsson K, Frykberg L. (1995) Cloning of Ligand-Binding Domains of Bacterial Receptors by Phage Display. *BioTechniques* **18**: 878-85

Jacobsson K, Frykberg L. (1996) Phage Display Shot-Gun Cloning of Ligand-Binding Domains of Prokaryotic Receptors Approaches 100% Correct Clones. *BioTechniques* **20**: 1070-1081

Jacobsson K, Rosander A, Bjerketorp J, Frykberg L. (2003) Shotgun Phage Display – Selection for Bacterial Receptors or other Exported Proteins. *Biol Proced Online* **5**: 123-135

Kaprelyants AS, Mukamolova GV, Kell DB. (1994) Estimation of dormant *Micrococcus luteus* cells by penicillin lysis and by resuscitation in cell-free spent culture medium at high dilution. *FEMS Microbiol Lett* **115**: 347-352

Keer J, Smeulders MJ, Williams HD. (2001) A *purF* mutant of *Mycobacterium smegmatis* has impaired survival during oxygen-starved stationary phase. *Microbiology* **147**: 473-481

Kell DB and Young M. (2000) Bacterial dormancy and culturability: the role of autocrine growth factors. *Current Opinion in Microbiology* **3**: 238-243

Kiefhaber T, Rudolph R, Kohler HH, Buchner J. (1991) Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation. *Biotechnology (N Y)*. **9**: 825-9

Lim A, Eleuterio M, Hutter B, Murugasu-Oei B, Dick T. (1999) Oxygen Depletion-Induced Dormancy in *Mycobacterium bovis* BCG. *J Bact* **181**: 2252-2256

Lockwood DNJ, Colston MMJ, Khanolkar-Young SR. (2002) The detection of *Mycobacterium leprae* protein and carbohydrate antigens in skin and nerve from leprosy patients with Type 1 (Reversal) reactions. *Am J Trop Med Hyg* **66**: 409-415

Loebel RO, Shorr E, Richardson HB. (1933) The influence of adverse conditions upon the respiratory metabolism and growth of human tubercle bacilli. *J Bact* **26**: 167-200

Lowrie DB, Tascon RE, Bonato VLD, Lima VM, Faccioli LH, Stavropoulos E, Colston MJ, Hewinson RG, Moelling K, Silva CL. (1999) Therapy of tuberculosis in mice by DNA vaccination. *Nature* **400**: 269-271

McCune RM, Tompsett R. (1957a) Fate of *Mycobacterium tuberculosis* in Mouse Tissues as Determined by the Microbial Enumeration Technique. I. The Persistence of Drug-Susceptible Tubercle Bacilli in the Tissues despite Prolonged Antimicrobial Therapy. *J Exp Med* **104**: 737-762

McCune RM, Tompsett R, McDermott W. (1957b) Fate of *Mycobacterium tuberculosis* in Mouse Tissues as Determined by the Microbial Enumeration

Technique. II. The Conversion of Tuberculous Infection to the Latent State by the Administration of Pyrazinamide and a Companion Drug. *J Exp Med* **104**: 763-802

McCune RM, Feldman FM, Lambert HP, McDermott W. (1966a) The capacity of tubercle bacilli to survive sterilization in mouse tissues. *J Exp Med* **123**:445-468

McCune RM, Feldman FM, Lambert HP, McDermott W. (1966b) The capacity of tubercle bacilli to survive sterilization in mouse tissues. *J Exp Med* **123**: 469-486

McKinney JD. (2000) *In vivo veritas*: the search for TB drug targets goes live. *Nature Medicine* **6**: 1330 –1333

McKinney JD, zu Bentrup KH, Munoz-Elias EJ, Micrak A, Chen B, Chan WT, Swenson D, Saccinetti JC, Jacobs WR, Russell DG. (2000) Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **406**: 735-738

Michele TM, Ko C, Bishai WR. (1999) Exposure to Antibiotics Induces Expression of the *Mycobacterium tuberculosis sigF* gene: Implications for Chemotherapy against Mycobacterial Persistors.

Misawa S, Kumagai I. (1999) Refolding of Therapeutic Proteins Produced in *Escherichia coli* as Inclusion bodies. *Biopolymers (Peptide Science)* **51**: 297-307

Mukamolova GV, Kaprelyants AS, Kell DB. (1995) Secretion of an antibacterial factor during resuscitation of dormant cells in *Micrococcus luteus* cultures held in an extended stationary phase. *Antonie Van Leeuwenhoek* **67**: 289-295

Mukamolova GV, Kell DB, Kaprelyants AS. (1999) Stimulation of the multiplication of *Micrococcus luteus* by an autocrine growth factor. *Arch Microbiol* **172**: 9-14

Mukamolova GV, Kaprelyants AS, Young DI, Young M, Kell DB. (1998) A bacterial cytokine. *Proc Natl Acad Sci USA* **95**: 8916-21

Mukamolova GV, Turapov OA, Kazarian K, Telkov M, Kaprelyants AS, Kell DB, Young M. (2002a) The *rpf* gene of *Micrococcus luteus* encodes an essential secreted growth factor. *Molecular Microbiology* **46** (3): 611-621

Mukamolova GV, Turapov OA, Young DI, Kaprelyants AS, Kell DB, Young M. (2002b) A family of autocrine growth factors in *Mycobacterium tuberculosis*. *Molecular Microbiology* **46** (3): 623-635

Mukamolova GV, Murzin AG, Salina EG, Demina GR, Kell DB, Kaprelyants AS, Young M. (2006) Muralytic activity of *Micrococcus luteus* Rpf and its relationship to physiological activity in promoting bacterial growth and resuscitation. *Molecular Microbiology* **59**: 84-98

Old RW, Primrose SB. (1994) Principles of gene manipulation: an introduction to genetic engineering – 5<sup>th</sup> edition. Blackwell Science, UK.

Opie EL, Aronson JD. (1927) Tubercle bacilli in latent tuberculous lesions and in lung tissue without tuberculous lesions. *Arch Pathol Lab Med* **4**: 1-21

Orme IM. (2001) The latent tuberculosis bacillus (I'll let you know if I ever meet one). *Int J Tuberc Lung Dis* **5**: 589-593

Orrell JM, Brett SJ, Ivanyi J, Coghill G, Grant A, Beck JS. (1991) Measurement of tissue distribution of immunoperoxidase staining with polyclonal anti-BCG serum in lung granulomata of mice infected with *Mycobacterium tuberculosis*. *J Pathol.* **164**: 41-45

Pai SR, Actor JK, Sepulveda E, Hunter RL, Jagannath C. (2000) Identification of viable and non-viable *Mycobacterium tuberculosis* in mouse organs by directed RT-PCR for antigen 85B mRNA. *Microbial Pathogenesis* **28**: 335-342

Parrish NM, Dick JD, Bishai WR. (1998) Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends in Microbiology* **6**: 107-112

Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S, Kriakov J, Lawrence JG, Jacobs WR Jr, Hendrix RW, Hatfull GF. (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**: 171-182

Prinz WA, Aslund F, Holmgren A, Beckwith, J. (1997) The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem.* **272**: 15661-67

Radhakrishnan VV, Mathai A, Radhakrishnan NS, Rout D, Sehgal S. (1991) Immunohistochemical demonstration of mycobacterial antigens in intracranial tuberculoma. *Indian J Exp Biol.* **29**: 641-644

Ravagnani A, Finan CL, Young M. (2005) A novel firmicute protein family related to the actinobacterial resuscitation promoting factors by non-orthologous domain displacement. *BMC Genomics* **6**: 39

Rickman L, Scott C, Hunt DM, Hutchinson T, Menendez MC, Whalan R, Hinds J, Colston MJ, Green J, Buxton RS. A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the *rpjA* gene coding for a resuscitation promoting factor. *Mol Microbiol.* **56**: 1274-1286

Rosander A, Bjerketorp J, Frykberg L, Jacobsson K. (2002) Phage display as a novel screening method to identify extracellular proteins. *J. Microbiol. Methods* **51**: 43-55

Schauder S & Bassler B. (2001) The languages of bacteria. *Genes and Development* **15**: 1468-1480

Schein, C.H. and Noteborn, M.H.M.(1988). Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature. *Bio/Technology* **6**: 291–294.

Schein CH. (1991) Optimizing protein folding to the native state in bacteria. *Curr Opin Biotechnol.* **2**(5):746-50.

Seiler P, Ulrichs T, Bandermann S, Pradl L, Jorg S, Krenn V, Morawietz L, Kaufmann SHE, Aichele P. (2003) Cell-Wall Alterations as an Attribute of *Mycobacterium tuberculosis* in Latent Infection. *J Inf Dis* **188**: 1326-1331

Shleeva MO, Bagramyan K, Telkov MV, Mukamolova G, Young M, Kell DB, Kaprelyants AS. (2002) Formation and resuscitation of 'non-culturable' cells of *Rhodococcus rhodochrous* and *Mycobacterium tuberculosis* in prolonged stationary phase. *Microbiol* **148**: 1581-91

Shleeva M, Mukamolova GV, Young M, Williams HD, Kaprelyants AS. (2004) Formation of 'non-culturable' cells of *Mycobacterium smegmatis* in stationary phase in response to growth under suboptimal conditions and their Rpf-mediated resuscitation. *Microbiology* **150**: 1687-1697

Smeulders MJ, Keer J, Speight RA, Williams HD. (1999) Adaptation of *Mycobacterium smegmatis* to Stationary Phase. *J Bact* **181**: 270-283

Smith GP. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**: 1315-17

Sun Z, Zhang Y. (1999) Spent Culture Supernatant of *Mycobacterium tuberculosis* H37Ra Improves Viability of Aged Cultures of This Strain and Allows Small Inocula To Initiate Growth. *J Bact* **181**: 7626-28

Swartz, JR. (2001) Advances in *Escherichia coli* production of therapeutic proteins. *Current Opinion in Biotechnology* **12**: 195-201

Tufariello JM, Jacobs WR Jr, Chan J. (2004) Individual *Mycobacterium tuberculosis* resuscitation-promoting factor homologues are dispensible for growth *in vitro* and *in vivo*. *Infect Immun* **72**: 515-526

Van den Bos IC, Khanolkar-Young DS, Das PK, Lockwood DN. (1999) Immunohistochemical detection of PGL-1, LAM, 30kD and 65 kD antigens in leprosy infected paraffin preserved skin and nerve sections. *Lepr Rev* **70**: 272-280

Verhagen C, Faber W, Klatser P, Buffing A, Naafs B, Das P. (1999) Immunohistological analysis of *in situ* expression of mycobacterial antigens in skin lesions of leprosy patients across the histopathological spectrum: association of mycobacterial lipoarabinomannan (LAM) and *Mycobacterium leprae* phenolic glycolipid-1 (PGL-1) with leprosy reactions. *Am J Path* **154**: 1793-1804

Wayne LG. (1976) Dynamics of submerged growth of *Mycobacterium tuberculosis* under aerobic and anaerobic conditions. *Am Rev Resp Dis* **114**: 807-811

Wayne LG. (1994) Dormancy of *Mycobacterium tuberculosis* and Latency of Disease. *Eur J Clin Microbiol Infect Dis* **13** (11): 908-914



Wayne LG & Lin K-Y. (1982) Glyoxylate metabolism and aaptation of *Mycobacterium tuberculosis* to survival under anaerobic conditions. *Infect. Immun.* **37**: 1042-1049

Wayne LG, Hayes LG. (1996). An *in vitro* Model for Sequential Study of Shiftdown of *Mycobacterium tuberculosis* through Two Stages of Nonreplicating Persistence. *Infect Immun* **64**: 2062-2069

Wayne LG, Hayes LG. (1998). Nitrate reduction as a marker for hypoxic shiftdown of *Mycobacterium tuberculosis*. *Tubercle Ling Dis.* **79**: 127-132

Willats WGT. (2002) Phage display: practicalities and prospects. *Plant Molecular Biology* **50**: 837-854

Wilson DS, Nock S. (2001) Functional Protein Microarrays. *Curr Opin Chem Biol* **6**: 81-85

de Wit D, Wootton M, Dhillon J, Mitchison DA. (1995) The bacterial content of mouse organs in the Cornell model of dormant tuberculosis. *Tubercle andLung Disease* **76**: 555-562

Yeremeev VV, Kondratieva TK, Rubakova EI, Petrovskaya SN, Kazarian KA, Telkov MV, Biketov SF, Kaprelyants AS, Apt AS. (2003) Proteins of the Rpf Family: Immune Cell Reactivity and Vaccination Efficacy against Tuberculosis in Mice. *Infect Immun* **71**: 4789-4794

Yuan Y, Crane DD, Simpson RM, Zhu YQ, Hickey MJ, Sherman DR, Barry CE. (1998) The 16-kDA alpha-crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages. *PNAS* **16**: 9678-9583

Zhang Y, Yang Y, Woods A, Cotter RJ, Sun Z. (2001) Resuscitation of Dormant *Mycobacterium tuberculosis* by Phospholipids or Specific Peptides. *Biochemical and Biophysical Research Communications* **284**: 542-547

Zhu W, Plikaytis BB, Shinnick TM. (2003) Resuscitation factors from mycobacteria: homologs of *Micrococcus luteus* proteins. *Tuberculosis* **83**: 261-269

## Appendices

- Appendix *i*    DNA sequences, amino acid sequences and composition analysis of *M. tuberculosis* RpfS
- Appendix *ii*    Media and solutions
- Appendix *iii*    Table of clones used
- Appendix *iv*    Vector maps
- Appendix *v*    MPN table
- Appendix *vi*    Table showing viable counts in *M bovis* BCG dormancy experiments (flask model) on solid agar and liquid media with and without Rpf
- Appendix *vii*    From luciferase to Lazarus – winning essay MRC Max Perutz prize 2003

## **APPENDIX *i***

### **DNA sequences, protein sequences and composition analysis of *M. tuberculosis* RpfS**

## DNA sequence of *M. tuberculosis* Rv0867 (*rpfA*)

```
>M. tuberculosis bacteria|Rv0867c|rpfA: 1221 bp -
                                     gaggg
attccgaccgcggatggaggcgggggacccaccaggtcc
gccgagatcggaccgagccgcatacggctccttggggg
gaagccaaacgtcgttcgccgatgtacgaccgcggcggg
tgatccctctccacggcgcaatccaacccgaacccgacagc
tgacctcgtaggcgtgtaccgagaggaattacctaacgt
1 - atg agt gga cgc cac cgt aag ccc acc aca
31 - tcc aac gtc agc gtc gcc aag atc gcc ttt
61 - acc ggc gca gta ctc ggt ggc ggc ggc atc
91 - gcc atg gcc gct cag gcg acc gcg gcc acc
121 - gac ggg gaa tgg gat cag gtg gcc cgc tgc
151 - gag tcg ggc gcc aac tgg tcg atc aac acc
181 - ggc aac ggt tac ctc ggt gcc ttg cag ttc
211 - act caa agc acc tgg gcc gca cat ggt ggc
241 - ggc gag ttc gcc ccg tcg gct cag ctg gcc
271 - agc cgg gag cag cag att gcc gtc ggt gag
301 - cgg gtg ctg gcc acc cag ggt cgc gcc gcc
331 - tgg ccg gtg tgc gcc cgc ggg tta tcg aac
361 - gca aca ccc cgc gaa gtg ctt ccc gct tcg
391 - gca gcg atg gac gct ccg ttg gac gcg gcc
421 - gcg gtc aac ggc gaa cca gca ccg ctg gcc
451 - ccg ccg ccc gcc gac ccg gcg cca ccc gtg
481 - gaa ctt gcc gct aac gac ctg ccc gca ccg
511 - ctg ggt gaa ccc ctc ccg gca gct ccc gcc
541 - gac ccg gca cca ccc gcc gac ctg gca cca
571 - ccc gcg ccc gcc gac gtc gcg cca ccc gtg
601 - gaa ctt gcc gta aac gac ctg ccc gca ccg
631 - ctg ggt gaa ccc ctc ccg gca gct ccc gcc
661 - gac ccg gca cca ccc gcc gac ctg gca cca
691 - ccc gcg ccc gcc gac ctg gcg cca ccc gcg
721 - ccc gcc gac ctg gcg cca ccc gcg ccc gcc
751 - gac ctg gca cca ccc gtg gaa ctt gcc gta
781 - aac gac ctg ccc gcg ccg ctg ggt gaa ccc
811 - ctc ccg gca gct ccc gcc gaa ctg gcg cca
841 - ccc gcc gat ctg gca ccc gcg tcc gcc gac
871 - ctg gcg cca ccc gcg ccc gcc gac ctg gcg
901 - cca ccc gcg ccc gcc gaa ctg gcg cca ccc
931 - gcg ccc gcc gac ctg gca cca ccc gct gcg
961 - gtg aac gag caa acc gcg ccg ggc gat cag
991 - ccc gcc aca gct cca ggc ggc ccg gtt ggc
1021 - ctt gcc acc gat ttg gaa ctc ccc gaa ccc
1051 - gac ccc caa cca gct gac gca ccg ccg ccc
1081 - ggc gac gtc acc gag gcg ccc gcc gaa acg
1111 - ccc caa gtc tcg aac atc gcc tat acg aag
1141 - aag ctg tgg cag gcg att cgg gcc cag gac
1171 - gtc tgc ggc aac gat gcg ctg gac tcg ctc
1201 - gca cag ccg tac gtc atc ggc
tgacqggctgagcgccgactttaaacgaaacccacccat
tcgtcggtaccgtcctcgaagaactggtgcttccacacg
ggaagccgcgccttgatggtctccaccaggtgcgcacag
gtgccaaacgccgccgccggtgatcgcgcgcaaccgcc
gccaccaggcgccctccccgacctgcaagacgcgcgatc
cggtg
```

## Amino acid sequence encoded by *M. tuberculosis* Rv0837 (*RpfA*)

```
>M. tuberculosis bacteria|Rv0867c|RpfA: 407 aa - POSSIBLE RESUSCITATION-PROMOTING
1 - MSGRHRKPTT SNVSVAKIAF TGAVLGGGGI AMAAQATAAT DGEWDQVARC ESGGNWSINT
61 - GNGYLGGLQF TQSTWAAHGG GEFAPSAQLA SREQQIAVGE RVLATQGRGA WPCVGRGLSN
121 - ATPREVLPA S AAMDAPLDAA AVNGEPAPLA PPPADPAPPV ELAANDLPAP LGEPLPAAPA
181 - DPAPPADLAP PAPADVAPPV ELAVNDLPAP LGEPLPAAPA DPAPPADLAP PAPADLAPPA
241 - PADLAPPAPA DLAPPVELAV NDLAPPLGEP LPAAPAE LAP PADLAPASAD LAPPAPADLA
301 - PPAPAE LAPP APADLAPPAA VNEQTAPGDQ PATAPGGPVG LATDLELPEP DPQPADAPPF
361 - GDVTEAPAET PQVSNIAYTK KIWQAIRAQD VCGNDALDSL AQPYYVIG
```

## Composition Analysis of Rv0867 RpfA of *M. tuberculosis*

Predicted Structural Class of the Whole Protein: R Deleage & Modification of Nishikawa & Ooi 1987

Analysis	Whole Protein
Molecular Weight	36751.00 m.w.
Length	374
1 microgram =	27.210 pMoles
Molar Extinction coefficient	32650 +/- 5%
1 A(280) =	1.13 mg/ml
Isoelectric Point	3.43
Charge at pH 7	-38.95

Amino Acid(s)	Number count	%by weight	%by frequency
Charged (RKHYCDE)	64	22.02	17.11
Acidic (DE)	48	15.79	12.83
Basic (KR)	9	3.67	2.41
Polar (NCQSTY)	55	16.76	14.71
Hvdrophobic (AILFWV)	156	38.34	41.71
AAla	89	17.22	23.80
CCys	3	0.84	0.80
DAsp	28	8.77	7.49
EGlu	20	7.03	5.35
F Phe	2	0.80	0.53
GGly	26	4.04	6.95
H His	1	0.37	0.27
I Ile	5	1.54	1.34
KLys	2	0.70	0.53
LLeu	37	11.40	9.89
M Met	1	0.36	0.27
NAsn	11	3.42	2.94
P Pro	78	20.61	20.86
QGln	15	5.23	4.01
RArg	7	2.98	1.87
S Ser	10	2.37	2.67
TThr	13	3.58	3.48
VVal	18	4.85	4.81
WTrp	5	2.53	1.34
YTyr	3	1.33	0.80
BAsx	0	0.00	0.00
ZGlx	0	0.00	0.00
XXx><	0	0.00	0.00
. Ter	0	0.00	0.00

## DNA sequence of *M. tuberculosis* Rv1009 (*rpfB*)

>*M. tuberculosis* bacteria|Rv1009|*rpfB*: 1081 aa - Probable resuscitation-promoting

```

gcccataacacccgcgcgcgcggagttgctcaacattgqdc
cggttcggttaccgctcttgatcgaacgggtggggcctc
taggttttcggaggggcccattttgctttttgttcgctgtg
taggtgggttgagtggttgcgcagggtcggggatatagcgcg
ttgactctacttaccaaaacttcacacagacccaatcacccg
1 - atg ttg cgc ctg gta gtc ggt gcg ctg ctg
31 - ctg gtg ttg gcg ttc gcc ggt ggc tat gcg
61 - gtc gcc gca tgc aaa acg gtg acg ttg acc
91 - gtc gac gga acc gcg atg cgg gtg acc acg
121 - atg aaa tcg cgg gtg atc gac atc gtc gaa
151 - gag aac ggg ttc tca gtc gac gac cgc gac
181 - gac ctg tat ccc gcg gcc ggc gtg cag gtc
211 - cat gac gcc gac acc atc gtg ctg cgg cgt
241 - agc cgt ccg ctg cag atc tcg ctg gat ggt
271 - cac gac gct aag cag gtg tgg acg acc gcg
301 - tcg acg gtg gac gag gcg ctg gcc caa ctc
331 - gcg atg acc gac acg gcg ccg gcc gcg gct
361 - tct cgc gcc agc cgc gtc ccg ctg tcc ggg
391 - atg gcg cta ccg gtc gtc agc gcc aag acg
421 - gtg cag ctc aac gac ggc ggg ttg gtg cgc
451 - acg gtg cac ttg ccg gcc ccc aat gtc gcg
481 - ggg ctg ctg agt gcg gcc ggc gtg ccg ctg
511 - ttg caa agc gac cac qtg gtg ccc gcc gcg
541 - acg gcc ccg atc gtc gaa ggc atg cag atc
571 - cag gtg acc cgc aat cgg atc aag aag gtc
601 - acc gag cgg ctg ccg ctg ccg ccg aac gcg
631 - cgt cgt gtc gag gac ccg gag atg aac atg
661 - agc cgg gag gtc gtc gaa gac ccg ggg gtt
691 - ccg ggg acc cag gat qtg acg ttc gcg gta
721 - gct gag gtc aac ggc gtc gag acc ggc cgt
751 - ttg ccc gtc gcc aac gtc gtg gtg acc ccg
781 - gcc cac gaa gcc gtg gtg cgg gtg gcc acc
811 - aag ccc ggt acc gag gtg ccc ccg gtg atc
841 - gac gga agc atc tgg gac gcg atc gcc gcc
871 - tgt gag gcc ggt ggc aac tgg gcg atc aac
901 - acc gcc aac ggg tat tac ggt ggt gtg cag
931 - ttt gac cag gcc acc tgg gag gcc aac gcc
961 - ggg ctg cgg tat gca ccc cgc gct gac ctc
991 - gcc acc cgc gaa gag cag atc gcc gtt gcc
1021 - gag gtg acc cga ctg cgt caa ggt tgg gcc
1051 - gcc tgg ccg gta tgt gct gca cga gcg ggt
1081 - gcg cgc
tgaccatccggetgctcgggcgcactgagatcaggcggc
tgcccaaagagctcgactttcggccgcgcgcaaatctctcg
gacagaacttcgtgcacgacgcgcaaacagggtgcgacggg
tggttgcgcctccggggtcagccgttcgcacctgglll
tggaggtcgggcgggctgggatcgetgacccctggcac
tgctc

```

## Amino acid sequence encoded by *M. tuberculosis* Rv1009 (*RpfB*)

>*M. tuberculosis* bacteria|Rv1009|*RpfB*: 362 aa - Probable resuscitation-promoting

```

1 - MLRLVVGALL LVLAFAGGYA VAACKTVTLT VDGAMRVTT MKSRVIDIVE ENGFSVDDRD
61 - DLYPAAGVQV HDADTIVLRR SRPLQISLDG HDAKQVWTTA STVDEALAQL AMTDTAPAAA
121 - SRASRVPLSG MALPVVSAKT VQLNDGGLVR TVHLPAPNVA GLLSAAGVPL LQSDHVVPA
181 - TAPIVEGMQI QVTRNRIKKV TERLPLPPNA RRVEDPEMMN SREVVLDPGV PGTQDVTFAV
241 - AEVNGVETGR LPVANVVVTP AHEAVVRVGT KPGTEVPPVI DGSIWDAIAG CEAGGNWAIN
301 - TGNGYGGVQ FDOGTWEANG GLRYAPRADL ATREEQIAVA EVTRLRQGWG AWPVCAARAG
361 - AR

```

## Composition Analysis of Rv1009 RpfB of *M. tuberculosis*

Predicted Structural Class of the Whole Protein: R Deleage & Roux Modification of Nishikawa & Ooi 1987

Analysis	Whole Protein		
Molecular Weight	35937.40 m.w.		
Length	341		
1 microgram =	27.826 pMoles		
Molar Extinction coefficient	39620 +/- 5%		
1 A(280) =	0.91 mg/ml		
Iso electric Point	5.19		
Charge at pH 7	-7.30		
Amino Acid(s)	Number count	%by weight	%by frequency
Charged (RKHYCDE)	82	30.71	24.05
Acidic (DE)	39	13.19	11.44
Basic (KR)	31	12.93	9.09
Polar (NCQSTY)	72	21.87	21.11
Hydrophobic (AILFWV)	133	36.36	39.00
A Ala	45	8.90	13.20
CCys	3	0.86	0.88
DAsp	21	6.73	6.16
EGlu	18	6.47	5.28
F Phe	3	1.23	0.88
G Gly	31	4.93	9.09
HHis	5	1.91	1.47
I Ile	12	3.78	3.52
KLys	7	2.50	2.05
L Leu	22	6.93	6.45
M Met	7	2.56	2.05
NAsn	12	3.81	3.52
P Pro	23	6.21	6.74
QGln	13	4.63	3.81
RArg	24	10.43	7.04
S Ser	13	3.15	3.81
TThr	27	7.60	7.92
VVal	45	12.41	13.20
WTrp	6	3.11	1.76
YTyr	4	1.82	1.17
BAsx	0	0.00	0.00
ZGlx	0	0.00	0.00
XXxx	0	0.00	0.00
. Ter	0	0.00	0.00



## DNA sequence of *M. tuberculosis* Rv1884 (*rpfC*)

```

M. tuberculosis H37Rv: Rv1884 (rpfC) (1228 bp) =
                                     gacgctc
cctgaataatcggatgctgtgcagatttggaatcactg
gagtttgctgtccgcgcctcctgcgcgcgcacaaactcga
ccgcgcacaaacgcgacatagtgcggtccgcgcacctcga
tagcctctatcaacgggcctgacgacagcaacacagtc
gtattgccagggcctaccgcgcgcctgatcaatccggtt
1 - gtg cat cct ttg ccg gcc gac cac ggc cgg
31 - tcg cgg tgc aat aga cac ccg atc tca cca
61 - ctc tct cta atc ggt aac gct tgc gcc act
91 - tcc ggc gat atg tgc agc atg aca aga atc
121 - gcc aag ccg ctc atc aag tcc gcc atg gcc
151 - gca gga ctc gtc acg gca tcc atg tgc ctc
181 - tcc acc gcc gtt gcc cac gcc ggt ccc agc
211 - ccg aac tgg gac gcc gtc gcg cag tgc gaa
241 - tcc ggg ggc aac tgg gcg gcc aac acc gga
271 - aac ggc aaa tac ggc gga ctg cag ttc aag
301 - ccg gcc acc tgg gcc gca ttc ggc ggt gtc
331 - ggc aac cca gca gct gcc tct cgg gaa caa
361 - caa atc gca gtt gcc aat cgg gtt ctc gcc
391 - gaa cag gga ttg gac gcg tgg ccg acg tgc
421 - ggc gcc gcc tct ggc ctt ccg atc gca ctg
451 - tgg tgc aaa ccc gcg cag gcc atc aag caa
481 - atc atc aac gag atc att tgg gca ggc att
511 - cag gca agt att ccg cgc
tgacggttggcggcgtgtgcggtctatgaccaggtcgac
gtatgtgtttggatcaggtcatggaaggttcggccacag
ttcacatggcagcgcgcgcgcgcgcacagatctggacattga
tcgcggatgtccgcaataccggccgggtctctgcgcggaaa
ccttcgaggcagagtggcttgacggtgccaccggccccc
cgctt

```

## Amino acid sequence encoded by *M. tuberculosis* Rv1884 (*RpfC*)

```

M. tuberculosis H37Rv: Rv1884 (rpfC) (1228 bp) = 408 amino acids
1 - VIELEPAEENK SRKRREIET LSLTGNAGWT SGDAETETL LKLTASAMA AGLLTAEEL
61 - STAVAHAGPS PNWDAVQCE NGENWAANTG NGKYGGGQPK PATWAAPGGV GNEAASRDL
121 - QIAVANRVLA EQGLDAWTC GARGSLPIAL WSKPAQGIKQ IINETIWAGT QASIPR

```

# Composition Analysis of Rv1884      RpfC of *M. tuberculosis*

Predicted Structural Class of the Whole Protein: Alpha  
Deleage & Roux Modification of Nishikawa & Ooi 1987

Analysis	Whole Protein
Molecular Weight	11141.10 m.w.
Length	109
1 microgram =	89.758 pMoles
Molar Extinction coefficient	35660 +/- 5%
1 A(280) =	0.31 mg/ml
Isoelectric Point	7.98
Charge at pH 7	0.85

Amino Acid(s)	Number count	%by weight	%by frequency
Charged (RKHYCDE)	16	18.83	14.68
Acidic (DE)	6	6.70	5.50
Basic (KR)	7	8.81	6.42
Polar (NCQSTY)	26	25.95	23.85
Hydrophobic (ALFWV)	46	43.22	42.20
A Ala	20	12.76	18.35
CCys	2	1.85	1.83
DAsp	2	2.07	1.83
EGlu	4	4.64	3.67
FPhe	2	2.64	1.83
GGly	16	8.20	14.68
HHis	0	0.00	0.00
I lie	9	9.14	8.26
KLys	4	4.60	3.67
LLeu	5	5.08	4.59
M Met	0	0.00	0.00
NAsn	7	7.17	6.42
P Pro.	8	6.97	7.34
QGln	7	8.05	6.42
RArg	3	4.21	2.75
SSer	6	4.69	5.50
TThr	3	2.72	2.75
VVal	4	3.56	3.67
WTrp	6	10.03	5.50
YTyr	1	1.46	0.92
BAsx	0	0.00	0.00
ZGlx	0	0.00	0.00
x Xxx	0	0.00	0.00
Ter	0	0.00	0.00

## DNA sequence of *M. tuberculosis* Rv2389 (*rpfD*)

*M. tuberculosis* bacteriophage Rv2389[*rpfD*]: 462 bp -  
gacaa  
tcgggtggccacccacgggacggtcgtggcgcggtcggg  
caaatcagccaccacgaacaggtagtggctcctgatgtt  
catcgaccagtcggtcagcaaccggaacagcccaacgac  
gcagatcgaccgcagtcggatcaaggtgatcatggacaa  
ggtgaacggccgctggctggccagcaaggtggagctgct  
1 - atg aca ccg ggt ttg ctt act act gcg ggt  
31 - gct ggc cga cca cgt gac agg tgc gcc agg  
61 - atc gta tgc acg gtg ttc atc gaa acc gcc  
91 - gtt gtc gcg acc atg ttt gtc gcg ttg ttg  
121 - ggt ctg tcc acc atc agc tcg aaa gcc gac  
151 - gac atc gat tgg gac gcc atc gcg caa tgc  
181 - gaa tcc ggc ggc aat tgg gcg gcc aac acc  
211 - ggt aac ggg tta tac ggt ggt ctg cag atc  
241 - agc cag gcg acg tgg gat tcc aac ggt ggt  
271 - gtc ggg tcg ccg gcg gcc gcg agt ccc cag  
301 - caa cag atc gag gtc gca gac aac att atg  
331 - aaa acc caa ggc ccg qgt gcg tgg ccg aaa  
361 - tgt agt tct tgt agt cag gga gac gca ccg  
391 - ctg ggc tcg ctc acc cac atc ctg acg ttc  
421 - ctc gcg gcc gag act gga ggt tgt tcg ggg  
451 - agc agg gac gat  
tgaggtgtgcctgcgttacattgctggccaccctagctg  
gtacgcgtatgcgatgatttggcatggatcttcggcag  
gcaccagttgagctgcccgccctggcgcccatgcccgga  
caaccgttcggggctctacctgcacgtcccgcttctgcctg  
accggttcgggtactgcgacttcaataacctacaccccg  
gccca

## Amino acid sequence encoded by Rv2389 (*RpfD*)

*M. tuberculosis* bacteriophage Rv2389[*rpfD*]: 154 aa - EVQVAKK PESUSCITAKLLKLVN TTHK  
1 - MIFGLHTTAG AGRTDRCAR LVETVFIETA VVAINFVALI GLSTNSKAE DINKATAAQ  
61 - ESGGNWAANT GNGLYGGLQI SQATWDSNGG VGSPAAASPQ QQIEVADNIM KTQGPQAWPK  
121 - CSSCSQGDAP LGSLLHILTF LAAETGGCSG SRDD

# Composition Analysis of Rv2389      RpfD of *M. tuberculosis*

Predicted Structural Class of the Whole Protein: Alpha  
Deleage & Roux Modification of Nishikawa & Doi 1987

Analysis	Whole Protein
Molecular Weight	10629.20 m.w.
Length	105
1 microgram =	94.080 pMoles
Molar Extinction coefficient	24520 +/- 5%
1 A(280) =	0.43 mg/ml
Isoelectric Point	3.66
Charge at pH 7	-9.04

Amino Acid(s)	Number count	%by weight	%by frequency
Charged (RKHYCDE)	21	23.98	20.00
Acidic (DE)	12	13.39	11.43
Basic (KR)	3	3.88	2.86
Polar (NCQSTY)	35	35.15	33.33
Hvdrophobic (AILFWV)	32	31.73	30.48
A Ala	13	8.70	12.38
C Cys	4	3.88	3.81
D Asp	9	9.75	8.57
G Glu	3	3.64	2.86
F Phe	1	1.38	0.95
G Gly	16	8.60	15.24
H His	1	1.29	0.95
I Ile	6	6.39	5.71
K Lys	2	2.41	1.90
L Leu	6	6.39	5.71
M Met	1	1.23	0.95
N Asn	5	5.37	4.76
P Pro	5	4.57	4.76
Q Gln	8	9.64	7.62
R Arg	1	1.47	0.95
S Ser	11	9.01	10.48
T Thr	6	5.71	5.71
V Val	2	1.86	1.90
W Trp	4	7.01	3.81
Y Tyr	1	1.54	0.95
B Asx	0	0.00	0.00
Z Glx	0	0.00	0.00
x Xxx	0	0.00	0.00
. Ter	0	0.00	0.00

### DNA sequence of *M. tuberculosis* Rv2450 (*rpfE*)

```
>M. tuberculosis bacteria|Rv2450c|rpfE: 516 bp -
                                     cgtgc
gagcggagccggggacggctgaatcggttgccgacccg
caggcggcggttagcagccgcgcgggcagatcaagcgc
tccgctgccgtgccggaccgagacggcacgtcccacgca
gcaccctcgaatgcacctggcagctaaccatgcccggtg
ggttggctttggcgcgcgcgcacccgcgaaaggaaacaacg
1 - ttg aag aac gcc cgt acg acg ctc atc gcc
31 - gcc gcg att gcc ggg acg ttg gtg acc acg
61 - tca cca gcc ggt atc gcc aat gcc gac gac
91 - gcg ggc ttg gac cca aac gcc gca gcc ggc
121 - ccg gat gcc gtg ggc ttt gac ccg aac ctg
151 - ccg ccg gcc ccg gac gct gca ccc gtc gat
181 - act ccg ccg gct ccg gag gac gcg ggc ttt
211 - gat ccc aac ctc ccc ccg ccg ctg gcc ccg
241 - gac ttc ctg tcc ccg cct gcg gag gaa gcg
271 - cct ccc gtg ccc gtg gcc tac agc gtg aac
301 - tgg gac gcg atc gcg cag tgc gag tcc ggt
331 - gga aac tgg tcg atc aac acc ggt aac ggt
361 - tac tac ggc ggc ctg cgg ttc acc gcc ggc
391 - acc tgg cgt gcc aac ggt ggc tcg ggg tcc
421 - gcg gcc aac gcg agc cgg gag gag cag atc
451 - cgg gtg gct gag aac gtg ctg cgt tcg cag
481 - ggt atc cgc gcc tgg ccg gtc tgc ggc cgc
511 - cgc ggc
tgaccaccgcgaaaaagacgcaaaaagctcccaaattccggt
cggattttgggagcttttgcgtctgctcgcgcaccccgga
gtagcgtcaaaccggtgaccgctacgccgcgtgaattcg
acatcgtgtgtgtacggggcgaccggcttcgtcggcaagc
tgaccgccgaataacctggcccgggcccgggggagatgcc
ggatc
```

### Amino acid sequence encoded by Rv2450 (*RpfE*)

```
>M. tuberculosis bacteria|Rv2450c|RpfE: 172 aa - PROBABLE RESUSCITATION-PROMOTING
1 - LKNARTTLIA AAIAGTLVTT SPAGIANADD AGLDPNAAAG PDAVGFDPNL PPAPDAAPVD
61 - TPPAPEDAGF DPNLPPPLAP DFLSPPAEAA PPVPVAYSVN WDAIAQCESG GNWSINTGNG
121 - YYGGLRFTAG TWRANGSGS AANASREEQI RVAENVLRSQ GIRAWPVCGR RG
```

## Composition Analysis of Rv2450 RpfE of *M. tuberculosis*

Predicted Structural Class of the Whole Protein: R Deleage &  
Roux Modification of Nishikawa & Ooi 1987

Analysis	Whole Protein
Molecular Weight	14904.20 m.w.
Length	145
1 microgram	67.095 pMoles
Molar Extinction coefficient	26840:1:5%
1 A(280)	0.56 mg/ml
Isoelectric Point	3.93
Charge at pH 7	-10.13

Amino Acid(s)	Number count	% by weight	%by frequency
Charged (RKHYCDE)	31	27.61	21.38
Acidic (DE)	18	14.56	12.41
Basic (KR)	8	8.38	5.52
Polar (NCQSTY)	30	22.29	20.69
Hydrophobic (AILFWV)	50	33.59	34.48
A Ala	23	10.97	15.86
CCys	2	1.38	1.38
DAsp	11	8.49	7.59
EGlu	7	6.06	4.83
FPh	4	3.95	2.76
GGly	17	6.51	11.72
HHis	0	0.00	0.00
I Ile	4	3.04	2.76
KLys	0	0.00	0.00
LLeu	7	5.32	4.83
M Met	0	0.00	0.00
NAsn	10	7.66	6.90
P Pro	21	13.68	14.48
QGln	3	2.58	2.07
RArg	8	8.38	5.52
SSer	8	4.68	5.52
TThr	4	2.71	2.76
VVal	8	5.32	5.52
WTrp	4	5.00	2.76
YTyr	3	3.28	2.07
BAsx	0	0.00	0.00
ZGlx	1	0.86	0.69
XXXX	0	0.00	0.00
. Ter	0	0.00	0.00

## **APPENDIX *ii***

### **Media and solutions.**

## Culture media

### *Sauton's medium/agar*

To make 1 litre:

0.5 g potassium dihydrogen orthophosphate

0.5 g Magnesium sulphate

4.0 g L-asparagine

6 ml glycerol

0.05 g ferric ammonium citrate

2.0 g tri-sodium citrate

0.1 ml 1% Zinc sulphate

Make up to 900ml with distilled water. To make agar, add 1g agar per 100ml.

Adjust pH to 7 with potassium hydroxide. Autoclave, cool, then add 2.5 ml 20% Tween-80 and 100 ml ADC.

### *Middlebrook agar*

Middlebrook 7H9 broth (Difco): prepared according to manufacturer's instructions. 1% agar added.

### *Luria-Bertani (LB) medium*

BactoTryptone            10 g/l

BactoYeast extract    5 g/l

NaCl                        10 g/l

For solid media add 1% agar. Adjust to pH 7.0 with sodium hydroxide, autoclave.

### *Nutrient Broth E*

Nutrient broth E (Lab M) 13g/l. For solid media add 1% agar. Autoclave.

### *Nutrient broth 2*

Nutrient Broth 2 (Oxoid) 25g/l, autoclave.

### *Nutrient Agar*

Nutrient Agar (Oxoid) 28g/l, autoclave.



*Lactate minimal media (LMM)*

Per 1 litre:

Potassium dihydrogen orthophosphate	4g
Ammonium chloride	4g
Magnesium sulphate	70mg
Ferrous sulphate	1 mg
Manganese chloride	0.5mg
Trace elements*	0.1ml
Thiamine	40mg
L-methionine	20mg
Biotin	5mg
Inosine	100mg

Adjust pH to 7.5 with sodium hydroxide, autoclave.

Before use, add 5g/l L-lactate and filter sterilise

\*Trace elements solution stored at 4°C:

per 10ml:

ZnSO <sub>4</sub> x 7H <sub>2</sub> O	5mg
H <sub>3</sub> BO <sub>3</sub>	5mg
CuSO <sub>4</sub> x 5H <sub>2</sub> O	2.4mg
Na <sub>2</sub> MoO <sub>4</sub>	2.5mg
H <sub>2</sub> SO <sub>4</sub>	0.05ml

*SOB*

Bacto Tryptone	2%
Bacto Yeast extract	0.5%
NaCl	10mM
KCl	2.5mM
MgCl <sub>2</sub>	10mM
MgSO <sub>4</sub>	10mM

Autoclave.

*Yeast-peptone induction medium*

15g select peptone

10g yeast extract

5g NaCl

Adjust pH to 7.2. Autoclave.

## **Buffers – DNA**

### *TAE buffer*

50x stock:

Tris base	242 g
EDTA	0.05 M
Glacial acetic acid	57.1 ml
H <sub>2</sub> O	To 1 l

### *Loading buffer*

Bromophenol blue	0.25%
Xylene cyanol	0.25%
Ficoll type 400 in H <sub>2</sub> O	57.1ml
EDTA	10mM

## **Buffers – protein**

### *PBS*

NaCl	8g
KCl	0.2g
Na <sub>2</sub> HPO <sub>4</sub>	2.88g
KH <sub>2</sub> PO <sub>4</sub>	0.48g

### *Charge buffer*

Nickel sulphate	50mM
-----------------	------

### *Binding buffer-with urea*

Urea (heat gently until dissolved)	8M
Tris HCl pH 8.0	40mM
NaCl	500 mM

### *Binding buffer – without urea*

Tris HCl pH 8.0	40mM
NaCl	500mM

#### *Elution buffer*

Imidazole	1M
NaCl	0.5M
Tris HCl pH8.0	40mM

#### *Strip buffer*

EDTA	100mM
NaCl	0.5M
Tris HCl pH 8.0	40mM

#### *Wash buffer*

Make up by adding 3ml elution buffer to 47ml binding buffer: contains 60mM imidazole

#### *Binding buffer – large-scale (2l) extractions*

Tris-HCl pH 8.0	25mM
NaCl	0.5M
10% glycerol	

#### *Elution buffer – large-scale (2l) extractions*

Imidazole	0.5M
NaCl	0.5M
Tris-HCl pH 8.0	25mM
10% glycerol	

#### *Sample buffer*

Tris-HCl	0.0625M
Glycerol	0.5%
SDS	2.3%
Bromophenol blue	1%

Freeze in 10ml aliquots. Add mercaptoethanol 2% before use

*Running buffer*

Tris base	0.025M
Glycerol	0.192M
SDS	0.1%

*Transfer buffer for immunoblotting*

Tris base	48mM
Glycine	39mM
Methanol	20%
pH 8.1-8.4	

*TBS*

Tris-HCl pH 8.0	10mM
NaCl	150mM

*Blocking buffer*

TBS containing 5% Marvel milk powder

*SDS-PAGE – 12.5% separating gel*

For 2 gels:

AA-MBA	3.275 ml
Tris-HCl pH 8.8	1.86 ml
H <sub>2</sub> O	2.725 ml
TEMED	15 µl
APS	10 µl

*SDS-PAGE – stacking gel*

For 2 gels:

AA-MBA	0.5 ml
Tris HCl pH 6.8	0.75 ml
H <sub>2</sub> O	1.7 ml
TEMED	15 µl
APS	15 µl

## APPENDIX *iii*

Table of clones referred to in the thesis

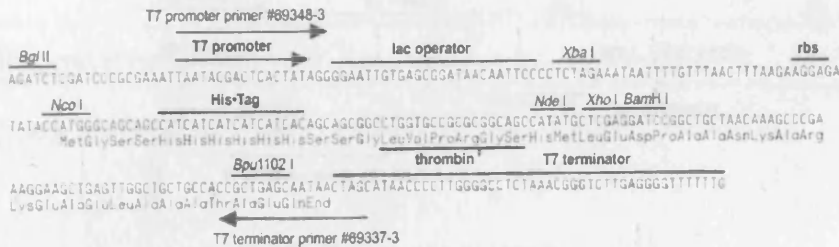
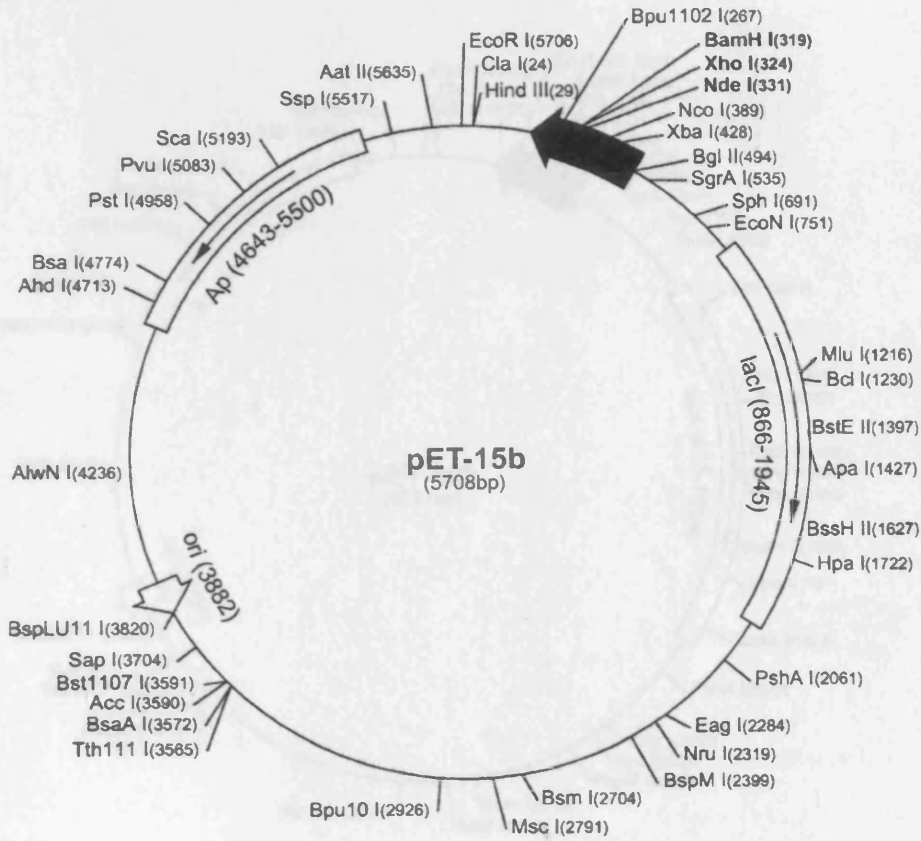
Clone	Insert	Vector	Host
AD1	<i>rpfD</i> <sup>†</sup>	TOPO®	DH5α
AD2	<i>rpfD</i> <sup>‡</sup>	TOPO®	DH5α
AD3	<i>rpfD</i>	pET-21a	DH5α
AD4	<i>rpfD</i>	pET-21a	Origami B (DE3) pLysS
AD5	<i>rpfB</i>	pET-19b	Origami B (DE3) pLysS
AD6	<i>rpfC</i>	pET-19b	Origami B (DE3) pLysS
AD7	<i>rpfA</i>	pET-19b	Origami B (DE3) pLysS
AD8	<i>rpfE</i>	pET-19b	Origami B (DE3) pLysS
AD9	<i>rpfE</i>	pET-15b.tev	DH5α
AD11	<i>rpfE</i>	pET-15b.tev	Origami B (DE3) pLysS
CA1261	<i>rpf</i> <sub>smeg</sub>	pSDY5	XL2-Blue
CA1272	-	pET-43.1b	XL2-Blue
CA1002	<i>rpfB</i>	pET-19b	HMS174
CA1318	<i>rpf</i> <sub>smeg</sub>	pET-43.1b	Origami (DE3)

† mutation at *Xho*I site

‡ no mutations

## **APPENDIX *iv***

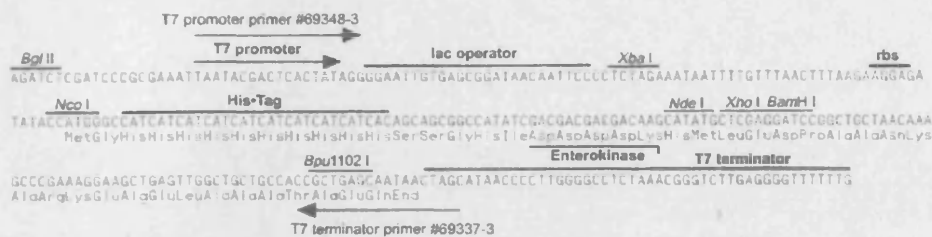
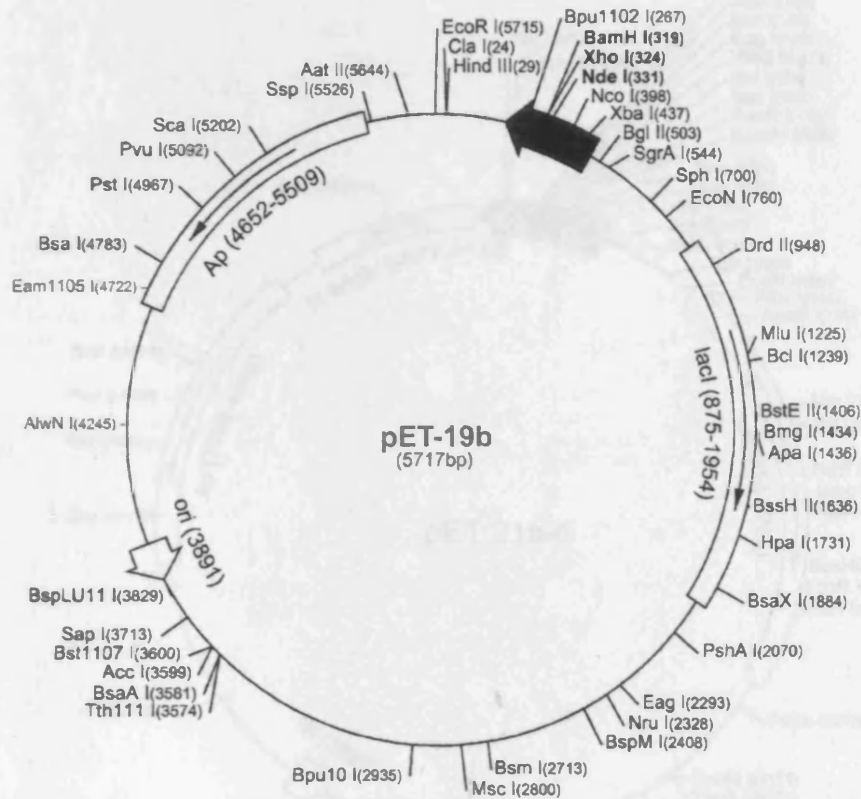
### **Vector maps**



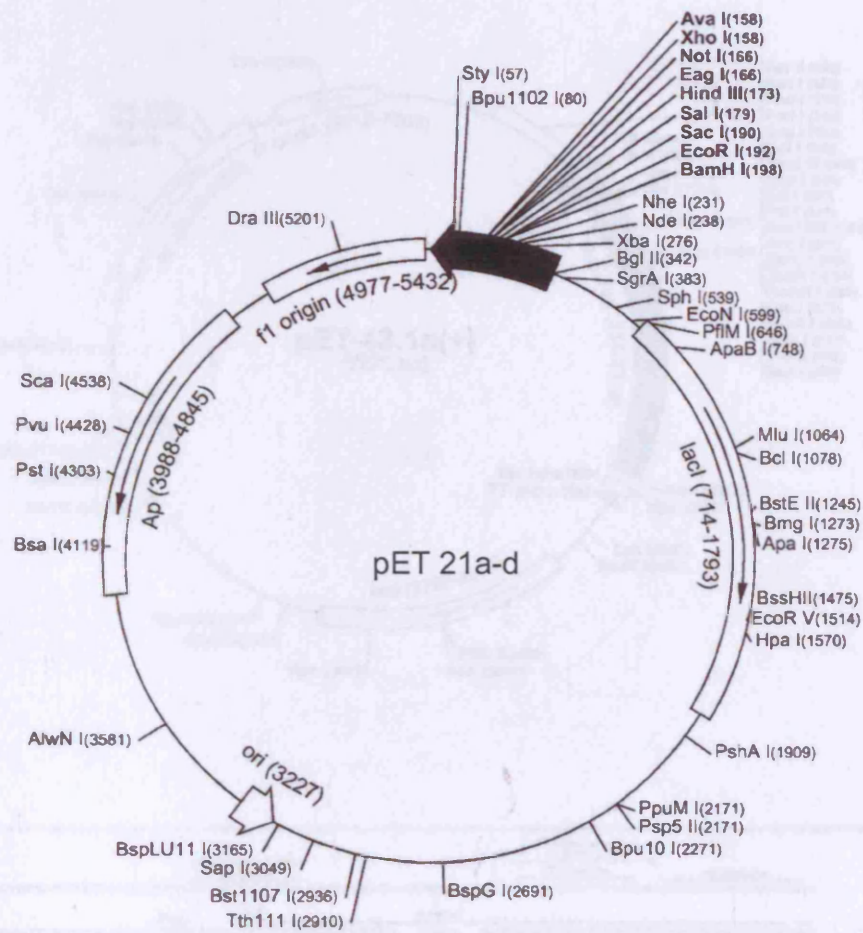
pET-15b cloning/expression region



## Vector map of pET 19.b (Novagen)



### Vector map of pET 21a-d (Novagen)

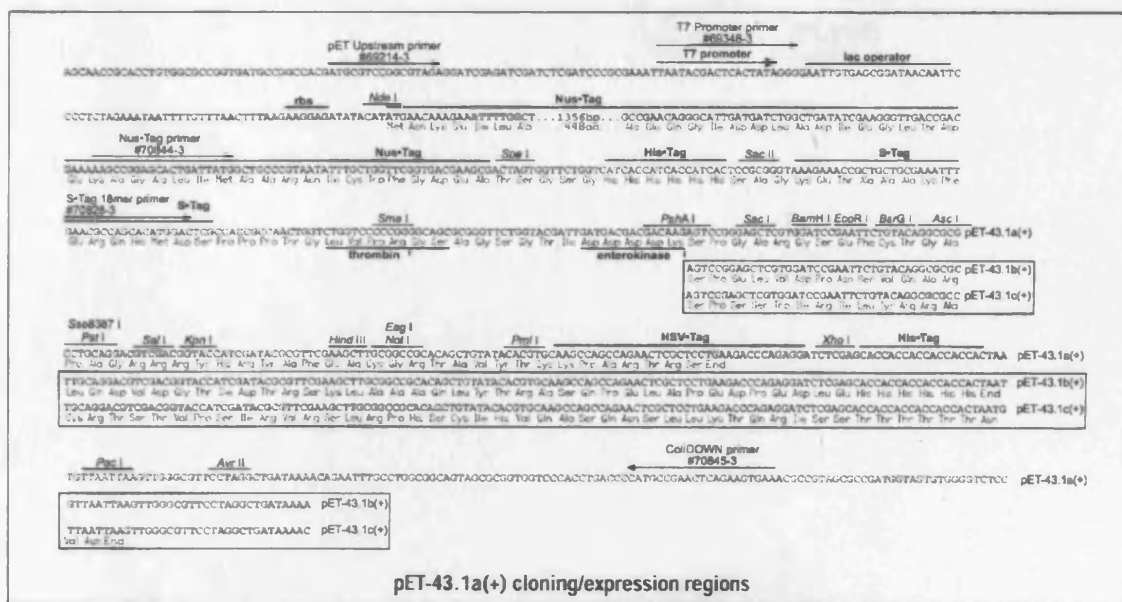
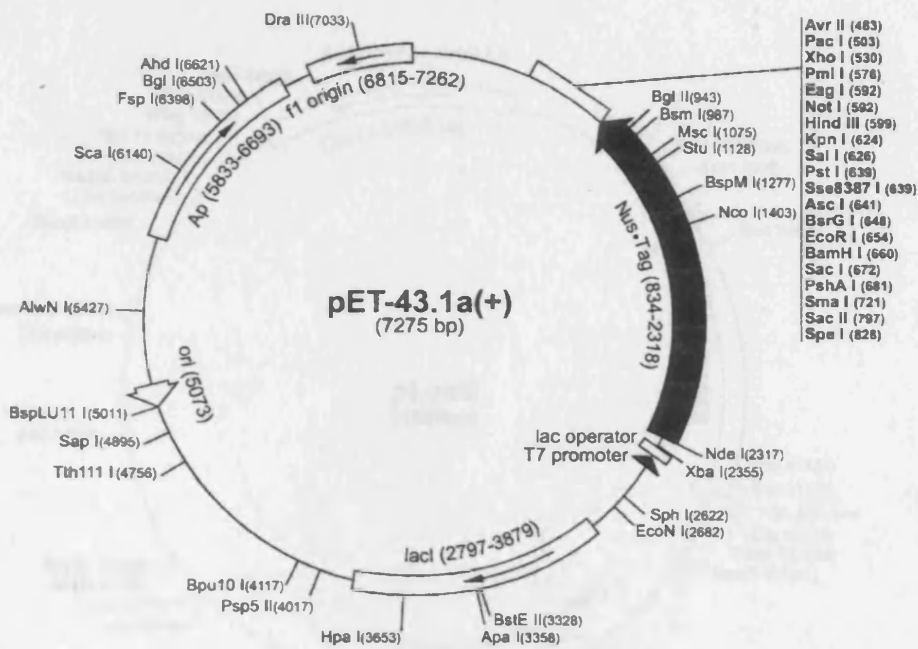


T7 promoter

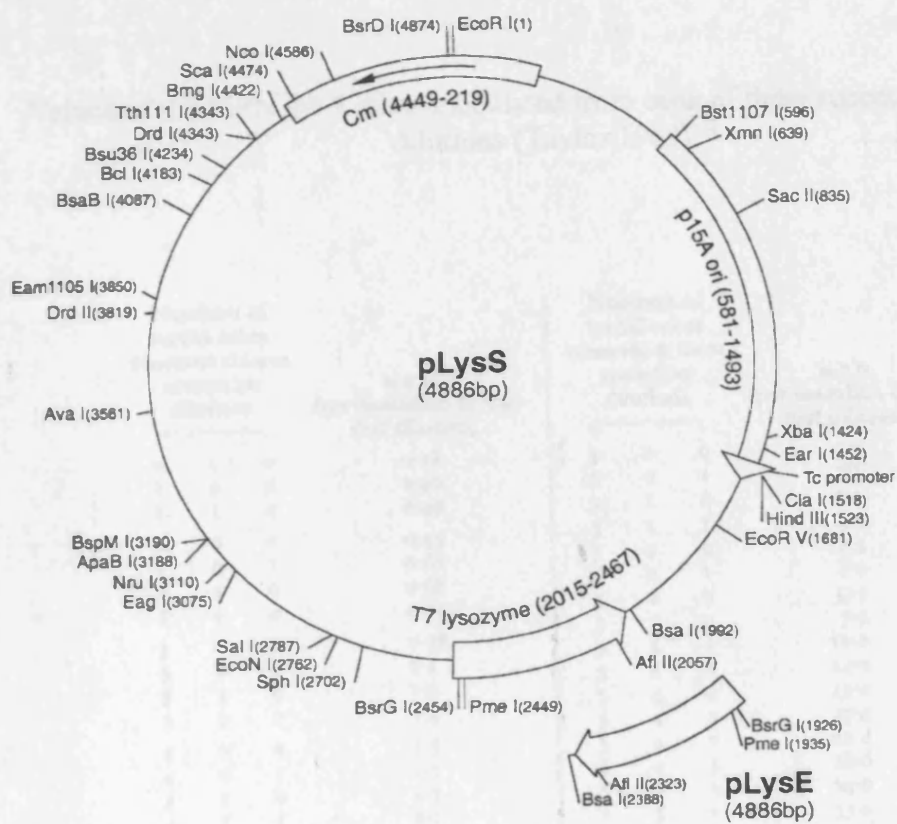
	pE1-21b	GGTGGGATCGAAATTCAGTTCCTGCAGAACTTGGGCGCACTCGAGCAACCAACGACCACCACTGA G...ArgAspProAsnSerSerSerValAsp...LeuAlaAlaIleGluGln...HisHisHisHisHisSerEnd
TACCATGGCTACG MetLaber	pE1-21c.d	GGTGGGATCGAAATTCAGCTCGCTGCAGAACTTGGGCGCACTCGAGCAACCAACCAACCAACCACTGA G...ArgAlaIleAlaAlaAlaProSerSerSerLeuArgProHisSerSerSerIleThrIleThrIleThrIleGlu

pET-21a-d(+) cloning/expression region

### Vector map of pET 43.1a-c (Novagen)



# Map of plasmid pLysS (Novagen)



## APPENDIX v

### MPN table

Values of the MPN for 5 tubes inoculated from each of three successive 10-fold dilutions (Taylor 1962)

Numbers of turbid tubes observed at three successive dilutions			M.P.N. (per inoculum of the first dilution)	Numbers of turbid tubes observed at three successive dilutions			M.P.N. (per inoculum of the first dilution)
0	1	0	0.18	5	0	0	2.3
1	0	0	0.20	5	0	1	3.1
1	1	0	0.40	5	1	0	3.3
2	0	0	0.45	5	1	1	4.6
2	0	1	0.68	5	2	0	4.9
2	1	0	0.68	5	2	1	7.0
2	2	0	0.93	5	2	2	9.5
3	0	0	0.78	5	3	0	7.9
3	0	1	1.1	5	3	1	11.0
3	1	0	1.1	5	3	2	14.0
3	2	0	1.4	5	4	0	13.0
4	0	0	1.3	5	4	1	17.0
4	0	1	1.7	5	4	2	22.0
4	1	0	1.7	5	4	3	28.0
4	1	1	2.1	5	5	0	24.0
4	2	0	2.2	5	5	1	35.0
4	2	1	2.6	5	5	2	54.0
4	3	0	2.7	5	5	3	92.0
				5	5	4	160.0

## APPENDIX vi

Results of viable counts from flask of *M bovis* BCG in Sauton's medium containing 200 ml of liquid culture with a headspace:broth ratio of 5:1 and constantly agitated

Age of culture (days)	Cfu/ml on agar	MPN in liquid medium	MPN with Rpf
0	20		
36	$10^6$		
60	$10^7$		
76	$6 \times 10^6$		
92	$2.6 \times 10^6$		
97	$10^6$		
109	$< 10^3^*$		
111	$< 10^4^*$	$1.2 \times 10^5$	$4.6 \times 10^5$
132	0	$> 8 \times 10^{5**}$	$> 8 \times 10^{5**}$
151	$3 \times 10^5$	$4.6 \times 10^6$	$1.2 \times 10^6$
154	$5 \times 10^5$		
165		$1.75 \times 10^6$	$8 \times 10^6$
168		$1.1 \times 10^6$	$8 \times 10^6$
176	$3.4 \times 10^5$		
183	$10^6$		

\*Maximum value – i.e. no growth was seen at the dilutions plated (not possible to ascertain the cfu/ml, as lower dilutions were not plated).

\*\*Minimum value – Growth occurred in all the MPN tubes

# Max Perutz Essay Prize 2003

## Winning essays

1st: **Angharad Davies**  
From Luciferase to Lazarus

2nd: **Chadrabala Shah**  
The Sting

3rd: **Babak Javid**  
Turning on the Heat

Special DNA prize: **Maxine Holder**  
Born to be wild type

# 1st st

## From Luciferase to Lazarus – a wake-up call for tuberculosis

Angharad Davies

My friend's father, Ken, had to strip naked and lie under a bright light for an hour a day (thirty minutes back, thirty minutes front) for two years. Worse, my great aunt as a young woman suffered the humiliation of months on what was then known as the 'venereal diseases' ward, being told she must have been a bad girl and that as punishment she wouldn't be able to have babies. In fact she, like Ken, had tuberculosis (TB) – pelvic in her case – in the days before it could be treated.

So feared was TB once that John Bunyan called it "the Captain of these men of death". Then, in the 1940s, drug treatments became available, and eventually, many believed, the disease would be eradicated.

In fact, today, one third of the world's population remains infected with the bacterium that causes TB – *Mycobacterium tuberculosis*. Nature has made a concession – not all those carrying the disease will actually become ill. Instead, the bacteria lie dormant in the body, evading the immune system, as if biding their time and stealthily waiting for the opportunity to attack when the time is right.

In a fortunate nine out of ten people, this never happens. In some cases, however, weakening of a person's immunity can cause the bacteria to reactivate – for instance, during chemotherapy for cancers, HIV infection, or simply old age. How this works, and what controls it, is not yet known.

Why not identify all the people who are carrying the TB bacterium in their bodies, and treat them with anti-tuberculosis drugs, before they become ill? Unfortunately this is not as easy as it sounds. The dormant bacteria are much less responsive to drugs than their wakeful counterparts. The chemical processes that go on inside all living organisms have been slowed right down – or even, some believe, stopped altogether. This means that the drugs, which act by interfering with these chemical processes, are unable to do their job. It's rather like trying to sabotage a car by putting sugar in the petrol tank – if the engine isn't running, nothing actually happens.

Similarly, the ability of these bacteria to become dormant accounts for the lengthy treatment needed for those who do become ill with TB. Even today, treatment takes a minimum of six months – a tall order for a well-funded and developed health system, let alone those in poorer parts of the world where TB is rife. It is one of the World Health Organisation's stated aims to find ways of decreasing the length of treatment for TB. To do this, a better understanding of the dormancy phenomenon is badly needed.

Recently, scientists at the University of Wales, Aberystwyth, have made a breakthrough in this field. Mike Young and his team have discovered small protein molecules, produced by



TB bacteria, which appear to switch the dormant bacteria back on. Mike raised the tantalising possibility that this might be a means of communication between the bacteria - if one of them senses that conditions are right to come out of the dormant state, it produces these proteins as a 'wake-up call' to alert the others - a sort of biological alarm clock.

The intriguing possibility of communication between bacteria is supported by evidence from the oceans. Certain aquatic bacteria can produce the protein luciferase, making them glow blue-green. This benefits their host, the squid, at night. In return for this illuminating activity, the squid allows the bacteria to live in a special light organ. Since glowing is such an energy-consuming form of rent, the bacteria only produce luciferase when they detect that there are enough of them around to make the effect worthwhile. This is known as 'quorum sensing'. So, whereas in the past scientists considered each single bacterium to be a completely separate individual, existing independently of all the others, it is now known that in fact they do produce signals to communicate with and influence one another. An analogy could be made with a colony of ants - each ant is a separate self-contained living organism but also behaves almost like a cell in the larger entity which is the colony.

Mike dubbed the TB 'wake-up' proteins 'Resuscitation Promoting Factors'. The aim of my research is to further investigate these resuscitation promoting factors, and to study whether by using them to 'wake up' dormant TB bacteria in the laboratory I can render the bacteria more susceptible to the action of anti-tuberculosis drugs. At first sight, resurrecting this particular Lazarus may seem the last thing that needs to be done. Nevertheless it may be that to hit the bacteria hard with drugs, we need to switch them back on first. The result could be shorter treatments and the curing of dormant disease. Ken says he's all for it.

### **Angharad Davies**

Department of Medical Microbiology  
Royal Free Hospital  
Pond Street  
London NW3

email: [a.davies@rfc.ucl.ac.uk](mailto:a.davies@rfc.ucl.ac.uk)  
telephone: }

MRC clinical research training Fellow since 22 October 2001